

A1

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
15. Februar 2001 (15.02.2001)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 01/11051 A2

- |                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(51) Internationale Patentklassifikation<sup>7</sup>: C12N 15/31, 15/54, 15/10, 15/62, 15/81, 9/12, 1/19, C07K 14/195, 16/18, 16/40, C12P 19/34, C12Q 1/68</p> <p>(21) Internationales Aktenzeichen: PCT/DE00/02657</p> <p>(22) Internationales Anmeldedatum: 7. August 2000 (07.08.2000)</p> <p>(25) Einreichungssprache: Deutsch</p> <p>(26) Veröffentlichungssprache: Deutsch</p> <p>(30) Angaben zur Priorität: 199 37 230.6 6. August 1999 (06.08.1999) DE</p> | <p>(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): LION BIOSCIENCE AG [DE/DE]; Im Neuenheimer Feld 517, D-69120 Heidelberg (DE).</p> <p>(72) Erfinder; und</p> <p>(75) Erfinder/Anmelder (nur für US): KILGER, Christian [DE/DE]; Handschuhsheimer Landstr. 4, D-69121 Heidelberg (DE). MOTZ, Michael [DE/DE]; Wundtstr. 9, D-69123 Heidelberg (DE). LÖSER, Eva [DE/DE]; Fritz-Frey-Str. 5/A37, D-69121 Heidelberg (DE). KÖGL, Manfred [DE/DE]; Hauptstr. 131/4, D-69214 Eppelheim (DE).</p> <p>(74) Anwalt: BOHMANN, Armin, K.; Bohmann &amp; Loosen, Pettenkoferstr. 35, D-80336 München (DE).</p> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

[Fortsetzung auf der nächsten Seite]

(54) Title: CHIMERIC PROTEINS

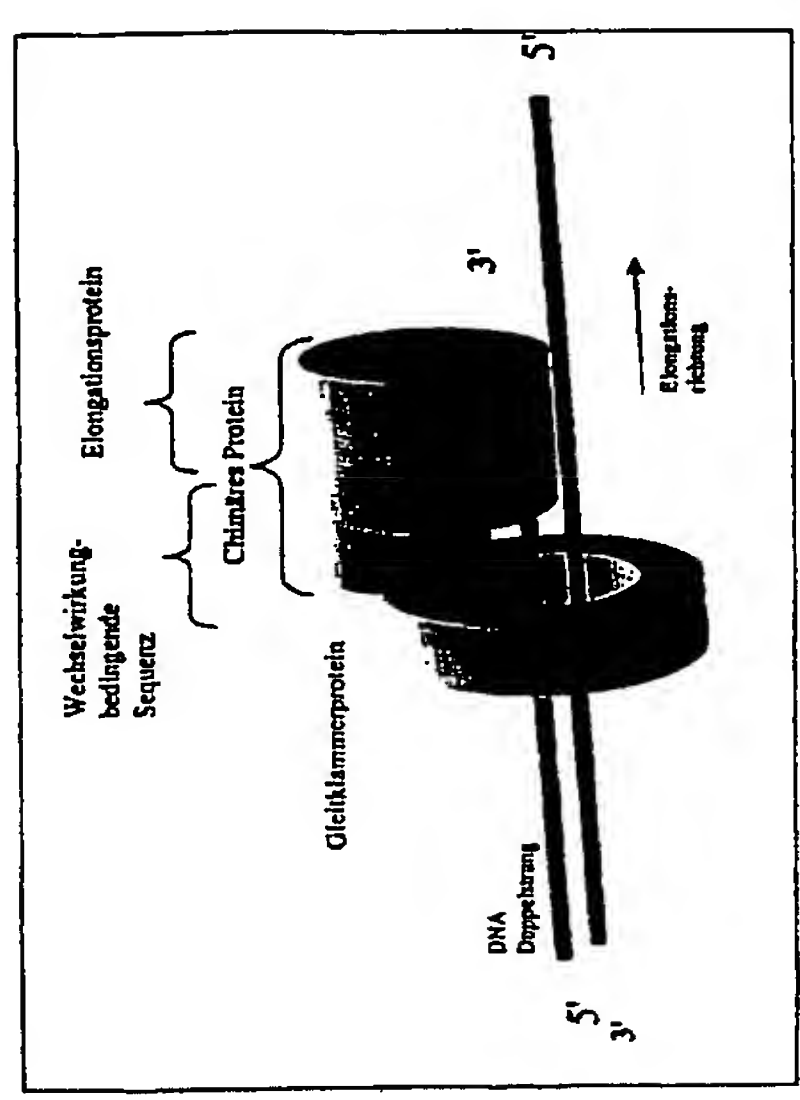
(54) Bezeichnung: CHIMÄRE PROTEINE

PCT/DE 00 / 0 2 6 5 7

DE 00/02657

2/13

Fig. 2



Wechsel... = Séquence établie par une interaction  
Elongationsproteine... = Protéines d'élongation  
Chimäres... = Protéine chimère  
Gleit... = Protéine d'anneau covalente  
DNA... = Double brin d'ADN  
Elongationsrichtung... = Sens d'élongation

(57) Abstract: The invention relates to a recombinant chimeric protein comprising a) a first domain with a nucleic acid synthesis activity and b) an interaction mediating sequence, whereby said interaction mediating sequence can form a complex through the nucleic acid synthesis activity and a slide bracketing protein. Said complex is different from the complex formed through the nucleic acid synthesis activity and/or the slide tying protein with their natural interaction partner(s).

(57) Zusammenfassung: Die Erfindung betrifft ein rekombinantes chimäres Protein, das a) eine erste Domäne mit Nukleinsäuresyntheseaktivität, und b) eine Wechselwirkung vermittelnde Sequenz umfaßt, wobei durch die Wechselwirkung vermittelnde Sequenz ein Komplex aus Nukleinsäuresyntheseaktivität und Gleitklammerprotein gebildet wird, der Komplex verschieden ist von dem Komplex, den die Nukleinsäuresyntheseaktivität und/or das Gleitklammerprotein mit ihren natürlichen Wechselwirkungspartner (n) ausbildet/ausbilden.

WO 01/11051 A2



(81) **Bestimmungsstaaten** (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Veröffentlicht:**

— Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

(84) **Bestimmungsstaaten** (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI,

Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

---

## Chimäre Proteine

---

10 Die vorliegende Erfindung betrifft rekombinante chimäre Proteine, die eine Nukleinsäuresyntheseaktivität aufweisen, diese enthaltende Komplexe, für diese codierende Nukleinsäuren, diese enthaltende Vektoren und Zellen, einen dagegen gerichteten Antikörper, Verwendungen dieser Proteine, diese enthaltende Kits sowie Verfahren zur Elongation, Amplifikation, reversen Transkription, Sequenzierung und Markierung von Nukleinsäuren.

Viele Proteine liegen in der Zelle nicht als Monomere vor, sondern sind Teil eines funktionellen multimeren Komplexes. Beispiele für solche Komplexe sind in fast allen Bereichen der Zellbiologie beschrieben (z.B. Transkription, Translation, Replikation, 20 Zytoskelett, Signaltransduktion, mRNA processing).

Die Wechselwirkung oder Bindung eines Proteins mit einem oder an ein anderes Protein, herein im folgenden Donorprotein und Akzeptorprotein genannt, erfolgt über bestimmte Aminosäuren oder Aminosäuresequenzen, die im gefalteten Protein meist an der Oberfläche des Proteins liegen und für die spezifische Bindung oder Wechselwirkung an oder mit dem Partner verantwortlich sind. Diese Aminosäuren bzw. eine solche Aminosäuresequenz wird herein im folgenden als "Wechselwirkungsbedingende Sequenz" oder Wechselwirkung-vermittelnde Sequenz bezeichnet. Den

Fachleuten ist dabei bekannt, daß die Aminosäuren, die an der Ausbildung einer Wechselwirkung beteiligt sind, nicht notwendigerweise in der primären Aminosäuresequenz unmittelbar aufeinanderfolgen, sondern ggf. mehr oder weniger stark konservierte Positionen innerhalb der Aminosäuresequenz des Proteins oder Peptids dafür verantwortlich zeichnen. Wenn hierin im folgenden von der Ermittlung der Wechselwirkung vermittelnden oder Wechselwirkung bedingenden Sequenz gesprochen wird, sollen damit beide vorstehend genannten Aspekte verstanden werden. In diesem Zusammenhang sind als Donorproteine solche Proteine zu verstehen, die mit einem weiteren Protein interagieren und gegebenenfalls ein weiteres Protein binden.

10 Gleichermaßen sind als Akzeptorproteine solche Proteine zu verstehen, die mit einem weiteren Protein interagieren und gegebenenfalls ein weiteres Protein binden

Es gibt eine Reihe von Methoden, um die Wechselwirkung-bedingende Sequenz einer Protein-Protein-Interaktionsstelle zu ermitteln. Dazu gehört (i) die Bestimmung der dreidimensionalen Struktur eines Komplexes aus Donor- und/oder Akzeptorprotein durch Röntgenstrukturanalyse oder (ii) NMR-Methoden. Eine weitere Methode ist (iii) eine Komplexbindung *in vitro* mit rekombinanten Proteinen zu rekonstituieren und durch gezielte Veränderung des Donor- oder Akzeptorproteins die Wechselwirkung-bedingende(n) Sequenz(en) zu ermitteln. Solche gezielten Veränderungen be-

20 inhalten die Mutation von einzelnen Aminosäuren, zum Beispiel zu Alanin (Alanin-scanning) oder die Deletion von Sequenzen im Protein. Führt die Veränderung oder Deletion einer Aminosäure oder Sequenz zum Verlust der Wechselwirkungs- oder Bindeaktivität, so ist sie Teil der Wechselwirkung-bedingende Sequenz, und umgekehrt, führt die Veränderung oder Deletion einer Aminosäure oder Sequenz nicht zum Verlust der Wechselwirkungs- oder Bindeaktivität so ist sie nicht Teil der Wechselwirkung-bedingende Sequenz. Auf diese Art und Weise läßt sich die Wechselwirkung-bedingende Sequenz definieren.

Eine weitere Methode zur Bestimmung der Wechselwirkung-bedingende Sequenz

30 eines Proteins beruht auf der Verwendung eines Zweihybridsystems, hierin im folgenden auch als "Y2H" abgekürzt. Y2H Systeme beruhen darauf, daß man ein Pro-



tein mit einer nachweisbaren Aktivität (wie z.B. das Enzym "Dihydrofolatreduktase") als zwei nicht kovalent miteinander verbundene Teile exprimiert. Dieses Protein ist inaktiv, wenn die beiden Teile sich nicht in räumlicher Nähe zu einander befinden. Die beiden zu untersuchenden Proteine, von denen man weiß, daß oder untersuchen möchte, ob sie miteinander in Wechselwirkung treten und somit einen Komplex ausbilden können, d.h. Donorprotein und Akzeptorprotein, werden unter Ausbildung jeweils eines Fusionsproteins mit jeweils einem der beiden Teile des Proteins mit der nachweisbaren Aktivität (wie z.B. Dihydrofolatreduktase) fusioniert und exprimiert so daß zwei Fusionsproteine entstehen. Wenn das fusionierte Donorprotein das fusionierte Akzeptorprotein bindet, kommen die beiden Hälften des nachweisbaren Proteins (wie z.B. Dihydrofolatreduktase) in räumliche Nähe zueinander. Dadurch wird die Aktivität des Proteins wiederhergestellt, die sodann nachgewiesen werden kann. Als Proteine mit nachweisbarer Aktivität können Enzyme (Dihydrofolatreduktase, beta-Galactosidase) Signaltransduktionsproteine (Cdc25 aus *Saccharomyces cerevisiae*) oder Transkriptionsaktivatoren (Gal4, LexA-VP16) eingesetzt werden. Die Ermittlung der Bindungsregion mit Hilfe des Zweihybridsystems beruht auf der gleichen Überlegung wie bei der *in-vitro*-Rekonstitution der Bindung: Führt die Veränderung oder Deletion einer Aminosäure oder Sequenz zum Verlust der Bindeaktivität so ist sie Teil der Wechselwirkung-bedingende Sequenz, und umgekehrt, führt die Veränderung oder Deletion einer Sequenz nicht zum Verlust der Bindeaktivität so ist sie nicht Teil der Wechselwirkung-bedingende Sequenz.

Zusätzlich dazu kann man die Wechselwirkung-bedingende Sequenz darüber definieren, daß eine Vielzahl von Fragmenten des Proteins bezüglich ihrer Interaktion untersucht und bestimmt wird, welche Teile des Proteins in den interagierenden oder wechselwirkenden Fragmenten immer vorhanden sind. Dieser stets vorhandene Bereich ist die Wechselwirkung-bedingende Sequenz. Um die in einer Wechselwirkung-bedingende Sequenz für die Wechselwirkung oder Bindung wesentlichen Aminosäuren zu identifizieren, kann man durch gezielte Mutationen einzelne Aminosäuren verändern. Verlust oder Erhöhung der Bindeaktivität weisen darauf hin, daß diese Positionen direkt an der Wechselwirkung beteiligt sind.

Unter den für in vitro Anwendungen besonders bevorzugten Komplexen gehören die thermostabilen Komplexe prokaryontischer und eukaryontischer Replikationsapparate, die als wichtige Enzymaktivität häufig Polymerasen beinhalten.

Wie eingangs erwähnt spielt die Wechselwirkung zwischen Proteinen *in vivo* eine große Rolle, so z. B. bei der Replikation der Nukleinsäure in biologischen Systemen. So sind hochprozessive Replikationsmechanismen bekannt, wobei es sich dabei zum einen um zelluläre Mechanismen und zum anderen um die bei den Bakteriophagen T4 und T7 auftretenden Replikationsmechanismen handelt.

Der Replikationsapparat umfaßt eine Vielzahl von Komponenten. Dazu gehören, unter anderem, a) Polymeraseaktivität aufweisende Proteine, b) Proteine, die an der Ausbildung einer Klammerstruktur beteiligt sind, wobei der Klammerstruktur unter anderem die Aufgabe zukommt, eine Polymeraseaktivität an ihr Template oder Matrize zu binden und die Bindung zu stabilisieren und somit die Dissoziationskonstante des Komplexes aus Polymerase und Nukleinsäure entsprechend zu ändern, c) Proteine, welche die Klammer auf das Template laden, d) Proteine, welche das Template stabilisieren und gegebenenfalls e) Proteine, welche die Polymerase an das Template führen.

Unter Polymeraseaktivität aufweisenden Proteinen werden hierin insbesondere solche Proteine verstanden, die ein oder mehrere Nukleotide oder Nukleoside an ein Nukleotid oder Nukleosid oder Polynukleotid oder Polynukleosid zu binden in der Lage sind. Dabei kann es sich in einem jeden der vorstehenden Fälle um Ribonukleotide/ Ribonukleoside oder Deoxynukleotide/ Deoxynukleoside oder Polymere davon handeln. Zu diesen Proteinen gehören somit neben DNA-Polymerasen auch RNA-Polymerasen, unabhängig davon, ob für die Polymerisationsreaktion des Proteins eine Matrize benötigt wird oder nicht. Diese Polymeraseaktivität aufweisende Proteine sind somit auch Nukleinsäuresyntheseaktivität aufweisende Proteine. Zu ihnen gehören auch die als solches im Stand der Technik bekannten Elongationsproteine.

Unter Elongationsprotein soll hierin im übrigen ein Polymeraseaktivität-aufweisendes Protein oder Komplex verstanden werden, das oder der mindestens eine oder mehrere der folgenden Eigenschaften aufweist: Verwendung von RNA als Template, Verwendung von DNA als Template, Synthese von RNA, Synthese von DNA, Exonukleaseaktivität in 5'-3' Richtung oder Exonukleaseaktivität in 3'-5' Richtung, Strangverdrängungsaktivität (engl. strand-displacement activity) und Prozessivität oder Nicht-Prozessivität.

- 10 DNA-Polymerasen gehören zu einer Gruppe von Enzymen, die einzelsträngige DNA als Template für die Synthese eines komplementären DNA Stranges verwenden. Diese Enzyme spielen eine bedeutende Rolle im Nukleinsäurestoffwechsel, einschließlich der Prozesse DNA-Replikation, Reparatur und Rekombination. DNA-Polymerasen wurden in allen zellulären Organismen identifiziert, von bakteriellen bis zu menschlichen Zellen, in vielen Viren sowie in Bakteriophagen (Kornberg, A. & Baker, T. A. (1991) DNA Replikation WH Freeman, New York, NY). Man faßt in der Regel die Archaeobakterien und die Eubakterien zusammen zu der Gruppe der Prokaryonten, der Organismen ohne echten Zellkern, und stellt ihnen die Eukaryonten, die Organismen mit echtem Zellkern, gegenüber. Gemeinsam sind vielen Polymerasen
- 20 aus den verschiedensten Organismen oftmals Ähnlichkeiten in der Aminosäuresequenz sowie Ähnlichkeiten in der Struktur (Wang, J., Sattar, A.K.M.A.; Wang, C.C., Karam, J.D., Konigsberg, W.H. & Steitz, T.A. (1997) Crystal Structure of pol  $\alpha$  family replication DNA polymerase from bacteriophage RB69. Cell 89, 1087-1099). Organismen wie der Mensch besitzen eine Vielzahl von DNA-abhängigen Polymerasen, von denen jedoch nicht alle für die DNA Replikation zuständig sind, sondern einige auch DNA-Reparatur durchführen. Replikative DNA-Polymerasen bestehen *in vivo* meist aus Proteinkomplexen mit mehreren Einheiten, welche die Chromosomen der zellulären Organismen und Viren replizieren. Eine generelle Eigenschaft dieser replizierenden Polymerasen ist im allgemeinen eine hohe Prozessivität, das heißt, deren
- 30 Fähigkeit, Tausende von Nukleotide zu polymerisieren ohne vom DNA Template ab-

zudissoziieren (Kornberg, A. & Baker, T. A. (1991) DNA Replikation. WH Freeman, New York, NY).

DNA-Polymerasen werden unter anderem durch zwei Eigenschaften charakterisiert, ihre Elongationsrate, das heißt die Anzahl der Nukleotide, die sie pro Sekunde in einen wachsenden DNA-Strang inkorporieren können und ihre Dissoziationskonstante. Wenn die Polymerase nach jedem Inkorporationsschritt eines Nukleotids in die wachsende Kette wieder vom Strang abdissoziiert (d.h. ein Elongationsschritt erfolgt pro Bindungsereignis), dann hat die Prozessivität den Wert 1 und die Polymerase ist nicht prozessiv.

- 10 Wenn die Polymerase für wiederholte Nukleinsäureinkorporationen mit dem Strang verbunden bleibt, dann wird die Elongation oder der Replikationsmodus und damit auch die Polymerase als prozessiv bezeichnet und kann einen Wert von mehreren Tausend erreichen (siehe hierzu auch: Methods in Enzymology Volume 262, DNA Replication, Edited by J. L. Campbell, Academic press 1995, pp. 270-280).

- 20 Die unter b) genannten Proteine bilden Strukturen aus, die entweder offen oder geschlossen sind, beispielsweise ringförmige oder halbringförmige Strukturen. Derlei Strukturen können durch eine oder mehrere Spezies von Proteinen ausgebildet werden. Dabei ist es möglich, daß eine der besagten Proteinspezies eine Polymeraseaktivität aufweist.

Die für die Ausbildung dieser Strukturen verantwortlichen Proteine werden, sofern sie keine Polymeraseaktivität aufweisen, hierin im folgenden als „Gleitklammerproteine“ oder „Klammerproteine“ bezeichnet.

- 30 Es ist bekannt, daß der Replikationsapparat in Archaea dem eukaryontischen Replikationsapparat ähnlich ist, obwohl die Genomorganisation in Eukaryonten und Archaea gänzlich verschieden ist und die zelluläre Struktur der Eubakterien der der Archaea ähnelt. (Edgell, D.R. and Doolittle, W.F. (1997). Archaea and the origin(s) of DNA replication proteins. Cell 89, 995-998.).

Die Gleitklammer ist häufig über ein oder mehrere weitere Proteine an ein Elongationsprotein gebunden, mit anderen Worten an das Elongationsprotein gekoppelt. Ein derartiges koppelndes Protein wird hierin im folgenden als Kopplungsprotein bezeichnet, wobei gegebenenfalls die Kopplung über eine Mehrzahl von Kopplungsproteinen erfolgen kann.

Die dreidimensionale Struktur von verschiedenen Gleitklammerproteinen wurde bereits bestimmt. Die Gesamtstruktur dieser Gleitklammern ist sehr ähnlich; die Bilder der ringförmig ausgebildeten Proteingesamtstruktur von PCNA, der  $\beta$ -Untereinheit und gp45 Ringe sind übereinandergelegt deckungsgleich (Kelman, Z. & O'Donnell, M. (1995) Structural and functional similarities of prokaryotic and eukaryotic sliding clamps. Nucleic Acids Res. 23, 3613-3620). Jeder Ring hat vergleichbare Dimensionen und eine zentrale Öffnung, die groß genug ist, um Duplex-DNA, also einen DNA-Doppelstrang, bestehend aus den zwei komplementären DNA-Strängen, zu umschließen.

Die Gleitklammer kann sich *in vivo* nicht selbst um die DNA herum positionieren, sondern muß um die DNA fixiert werden. In Prokaryonten und Eukaryonten besteht ein derartiger Proteinkomplex aus einer Vielzahl von Untereinheiten. Der Proteinkomplex erkennt das 3'-Ende des Primers eines „Primer-Template-Duplexes“, der zu einem längeren Doppelstrang durch Einbau von Nukleotiden ergänzt werden soll, und positioniert die Gleitklammer um die DNA.

Im Falle des Bacteriophagen T7 wird das gleiche Ziel, eine prozessive DNA-Synthese, wie unten definiert, mittels eines strukturell anderen Proteinkomplexes erreicht. Der Phage exprimiert eine eigene katalytische Polymerase, die T7-Polymerase, die das Genprodukt des *gene 5* und mit einem Protein aus dem Wirt *Escherichia coli*, dem Thioredoxin, eine Bindung eingeht und als Replikase eine hochprozessive DNA-Replikation ermöglicht. Auch hierbei kommt es zur Ausbildung einer Klammer, jedoch weist diese Klammer nicht die gleiche Struktur auf, wie z.B. im Falle des eukaryontischen PCNA.



Oft ist es nötig, wie zum Beispiel im Falle der humanen Polymerase  $\delta$ , daß Proteine (Kopplungsproteine) die Verbindung zwischen dem katalytisch aktiven Teil der Polymerase und einem Prozessivitätsfaktor schaffen. Ein Prozessivitätsfaktor ist eine Verbindung oder ein Molekül, das die Prozessivität einer Polymerase beeinflusst, bevorzugterweise erhöht. Ein Beispiel für einen Prozessivitätsfaktor stellen die Gleitklammerproteine dar. Beim Menschen ist dieses Kopplungsprotein die kleine Untereinheit der  $\delta$ -Polymerase (Zhang, S.-J., Zeng, X.-R., Zhang, P., Toomey, N.L., Chuang, R.Y., Chang, L.-S., and Lee, M.Y.W.T. (1994). A conserved region in the amino terminus of DNA polymerase  $\delta$  is involved in proliferating cell nuclear antigen binding. J. Biol. Chem. 270, 7988-7992). Im Falle der T7 Polymerase jedoch bindet der Prozessivitätsfaktor die katalytische Einheit der Polymerase direkt ohne Beteiligung eines Kopplungsproteins.

Die *in vitro* Anwendung von Proteinen mit Nukleinsäuresyntheseaktivität wie Polymerasen oder Elongationsproteinen ist in der Technik weit verbreitet, so z.B. bei der Polymerase-Kettenreaktion (PCR), Sequenzierung von Nukleinsäuren oder reversen Transkription. Dabei ist bei den meisten *in vitro* Anwendungen, wie PCR oder Sequenzierungsprozesse, Prozessivität eine wünschenswerte Eigenschaft, die allerdings die bislang in diesen Reaktionen eingesetzten thermostabilen Enzyme nach dem Stand der Technik nur in geringem Maße besitzen.

Die U.S. Patente 4,683,195, 4,800,195 und 4,683,202 beschreiben die Anwendung solcher thermostabiler DNA-Polymerasen in der Polymerase-Kettenreaktion (PCR). In der PCR wird unter Verwendung von Primern, Template (auch Matrize genannt), Nukleotiden, einer DNA-Polymerase, eines entsprechenden Puffers und bei geeigneten Reaktionsbedingungen DNA neu synthetisiert. In dieser PCR wird bevorzugt eine thermostabile Polymerase verwendet, die das zyklische thermische Aufschmelzen der DNA-Stränge übersteht. So wird häufig *Taq* DNA-Polymerase verwendet (U.S. Patent 4,965,188). Die Prozessivität der *Taq* DNA-Polymerase ist jedoch, wie oben ausgeführt, relativ gering im Vergleich zu derjenigen der T7-Polymerase.



DNA-Polymerasen finden auch bei der DNA -Sequenzbestimmung Anwendung (Sanger et al., *Proc. Natl. Acad. Sci.*, USA 74:5463-5467 (1997)). Eine der hierbei verwendeten Polymerasen kann z.B. die oben erwähnte *Taq*-Polymerase (U.S. Patent 5,075, 216) oder die Polymerase von *Thermotoga neapolitana* (WO 96/10640) oder andere thermostabile Polymerasen sein. Neuere Verfahren koppeln die exponentielle Amplifikation und die Sequenzierung eines DNA Fragmentes in einem Schritt, so daß es möglich ist, genomische DNA direkt zu sequenzieren. Eines der Verfahren, das sogenannte DEXAS-Verfahren (Nucleic Acids Res 1997 May 15;25(10):2032-2034 Direct DNA sequence determination from total genomic DNA. Kilger C, Pääbo S, Biol. Chem. 1997 Feb; 378(2):99-105 Direct exponential amplification and sequencing (DEXAS) of genomic DNA. Kilger C, Pääbo S und DE 19653439.9 sowie DE 19653494.1) verwendet eine Polymerase mit verminderter Diskriminierungsfähigkeit gegenüber Dideoxynukleotiden (ddNTPs) im Vergleich zu Deoxynukleotiden (dNTPs) sowie einen Reaktionspuffer, zwei Primer, die bevorzugt nicht äquimolar vorliegen, und die oben genannten Nukleotide, um dann in mehreren Zyklen eine komplette, sequenzspezifische DNA-Leiter eines Fragmentes zu erhalten, welches von den Primern flankiert ist. Eine Weiterentwicklung dieses Verfahrens besteht in der Verwendung eines Polymerasegemisches, wobei eine der beiden Polymerase zwischen ddNTPs und dNTPs diskriminiert, während die zweite eine verminderte Diskriminierungsfähigkeit aufweist (Nucleic Acids Res 1997 May 15;25(10):2032-2034 Direct DNA sequence determination from total genomic DNA. Kilger C, Pääbo S).

DNA-Polymerasen finden des weiteren Anwendung bei der reversen Transkription von RNA in DNA. Hierbei dient RNA als Template und die Polymerase synthetisiert einen komplementären DNA-Strang. Zur Anwendung kommt hier z.B. die thermostabile DNA-Polymerase aus dem Organismus *Thermus thermophilus* (Tth) (U.S. Patent 5,322,770).

Manche Polymerasen besitzen eine 'proof-reading'- Aktivität, also eine 3'-5' Exonukleaseaktivität. Diese Eigenschaft ist insbesondere dann wünschenswert, wenn das zu

synthetisierende Produkt mit einer niedrigen Fehlerrate bei der Nukleotidinkorporation hergestellt werden soll. Ein Beispiel stellen Polymerasen aus dem Organismus *Pyrococcus wosei* dar.

Die oben genannten Elongationsproteine, die in den vorstehend genannten Anwendungen zum Einsatz kommen, gehören *in vivo* größtenteils nicht zu den eigentlichen Replikationsenzymen, sondern es sind zumeist Enzyme, von denen man annimmt, daß sie bei der DNA-Reparatur beteiligt sind, weshalb deren Prozessivität relativ gering ist. Zudem besitzt jeder Organismus eine Vielzahl von Polymerasen die eine Vielzahl von

10    Eigenschaften besitzen.

Solche Elongationsproteine weisen, wie oben ausgeführt, z.B. die folgenden Eigenschaften auf: Verwendung von RNA als Template, Verwendung von DNA als Template, Synthese von RNA, Synthese von DNA, Exonukleaseaktivität in 5'-3' Richtung und Exonukleaseaktivität in 3'-5' Richtung, Strangverdrängungsaktivität (strand-displacement activity), Prozessivität oder Nicht-Prozessivität oder Thermostabilität oder Thermosensitivität auf. *In vivo* ist es jedoch häufig so, daß ein Proteinkomplex eine oder mehrere dieser Eigenschaften vereint. Insbesondere liegen *in vivo* oftmals Replikationskomplexe vor, deren Prozessivität, wie oben ausgeführt, durch das Vorhandensein eines Gleitklammerproteins gesteigert wird.

20   

Neben den oben genannten Bestandteilen von Replikationsapparaten findet man *in vivo* häufig auch andere DNA-modifizierende Proteine in größeren Komplexen mit weiteren Proteinen. In solchen Komplexen kommt es oftmals dazu, daß von einem ersten Protein "eine DNA-modifizierende Aktivität", wie beispielsweise terminale-Transkriptase-Aktivität, vereinigt wird mit einer oder mehreren, durch ein weiteres Protein in den Komplex eingebrachten Aktivität(en), wie beispielsweise Exonukleaseaktivität. Unter DNA-modifizierender Aktivität ist hierin jede enzymatische Aktivität, die zu einer chemischen, physikalischen oder strukturellen Veränderung einer Ausgangs-

30    Nukleinsäure führt, zu verstehen. Es kommt auch vor, daß eine DNA-modifizierende Aktivität erst durch die Interaktion mit mindestens einem weiteren Protein zustande

kommt. Weiter ist es möglich, daß eine DNA-modifizierende Aktivität verringert oder gesteigert wird durch die Interaktion mit mindestens einem weiteren Protein. Somit entsteht *in vivo* oftmals ein Proteinkomplex der z.B. die Summe der Einzelaktivitäten trägt oder dessen Aktivität gegenüber der Einzelaktivität verbessert ist.

Wie sich aus dem Vorstehenden ergibt, weisen *in vivo* Komplexe, die eine Nukleinsäuresyntheseaktivität umfassen, oftmals weitere, für die technische, d.h. *in vitro* Anwendung erwünschte Eigenschaften auf, die über die eigentliche Nukleinsäuresyntheseaktivität hinausgehen und von den weiteren, den Komplex ausbildenden Komponenten  
10 beigesteuert werden.

Eine unmittelbare technische Verwendung derartiger *in vivo* Komplexe z. B. zu Zwecken der Durchführung einer Sequenzierung von DNA, Durchführen einer Polymeraseketten-Reaktion oder Einführen von Markierungen in Nukleinsäuren scheiterte bisher aus einer Reihe von Gründen. Ein Grund war und ist die fehlende Kenntnis aller an der Ausbildung des interessierenden Komplexes beteiligter Faktoren oder Einzelkomponenten. Ein weiterer Grund besteht darin, daß der mehrere Komponenten umfassende Komplex neben den erwünschten auch eine oder mehrere unerwünschte Eigenschaften aufweist.

20 Ein zur *in vitro* Verwendung von *in vivo* Komplexen alternativer Ansatz sieht vor, Komponenten mit verschiedenem Ursprung, die jedoch für sich die erwünschte Eigenschaft besitzen, miteinander zu kombinieren und auf diese Weise den Komplex mit den erwünschten Eigenschaften auszubilden. Eine derartige Eigenschaft kann bspw. eine höhere Prozessivität sein. Dieser Ansatz scheiterte in der Technik daran, daß die den Komplex ausbildenden Einzelkomponenten, die miteinander in Wechselwirkung treten müßten, dies nicht oder nur schlecht taten und somit dem Komplex ebenfalls nicht die erwünschten Eigenschaften verliehen werden konnten.

30 Der vorliegenden Erfindung liegt somit die Aufgabe zugrunde, Proteine mit einer Nukleinsäuresyntheseaktivität bereitzustellen, die eine erhöhte Prozessivität aufweisen.

Es ist weiterhin eine Aufgabe der vorliegenden Erfindung, ein Verfahren bereitzustellen, das die Konstruktion derartiger Proteine erlaubt.

Schließlich ist es eine Aufgabe der Erfindung, Verfahren zur Amplifikation, insbesondere durch PCR, sowie Sequenzierung von Nukleinsäuren bereitzustellen, die die Amplifizierung und/oder Sequenzierung längerer Nukleinsäuren erlauben.

Die Aufgabe wird erfindungsgemäß gelöst durch ein rekombinantes chimäres Protein umfassend

- a) eine erste Domäne mit Nucleinsäuresyntheseaktivität, und
- b) eine Wechselwirkung vermittelnde Sequenz, wobei

durch die Wechselwirkung vermittelnde Sequenz ein Komplex aus Nukleinsäuresyntheseaktivität und Gleitklammerprotein gebildet wird und der Komplex verschieden ist von dem Komplex, den die Nukleinsäuresyntheseaktivität und/oder das Gleitklammerprotein mit ihren natürlichen Wechselwirkungspartner (n) ausbildet.

- 20 In einem weiteren Aspekt der Erfindung wird die Aufgabe gelöst durch ein chimäres Protein umfassend einen Nukleinsäuresyntheseaktivität aufweisenden Anteil, wobei der Anteil aus einem einen Nukleinsäuresyntheseaktivität aufweisenden Anteil und einen Wechselwirkung vermittelnden Anteil umfassenden Basisprotein stammt, und mindestens einen eine Wechselwirkung vermittelnden Anteil, wobei der Wechselwirkung vermittelnde Anteil verschieden ist von dem Wechselwirkung vermittelnden Anteil des Basisproteins.

In einer Ausführungsform ist dabei vorgesehen, daß das Protein den Wechselwirkung vermittelnden Anteil des Basisproteins umfaßt.

30

In einem weiteren Aspekt wird die Aufgabe gelöst durch ein chimäres Protein umfassend einen Nukleinsäuresyntheseaktivität aufweisenden Anteil, wobei der Anteil aus

einem Basisprotein stammt, das einen Nukleinsäuresyntheseaktivität aufweisenden Anteil aber keinen Wechselwirkung vermittelnden Anteil umfaßt, und einen Wechselwirkung vermittelnden Anteil.

Bei den erfindungsgemäßen chimären Proteinen kann vorgesehen sein, daß der Wechselwirkung vermittelnde Anteil eine Bindung zwischen der Nukleinsäuresyntheseaktivität und einem die Syntheseleistung der Nukleinsäuresyntheseaktivität beeinflussenden Faktor vermittelt. In einer bevorzugten Ausführungsform ist dabei vorgesehen, daß der Faktor ein Gleitklammerprotein ist.

10

In einer Ausführungsform kann vorgesehen sein, daß die Nukleinsäuresyntheseaktivität eine Konsensuspeptidsequenzen umfaßt, wobei die Sequenz ausgewählt ist aus der Gruppe, die die Sequenzen SEQ ID NO.: 1, 2 und 3 umfaßt.

In einer weiteren Ausführungsform kann vorgesehen sein, daß das Gleitklammerprotein eine Konsensuspeptidsequenz umfaßt, wobei die Sequenz ausgewählt ist aus der Gruppe, die die Sequenzen SEQ ID NO.: 4, 5, 6 und 7 umfaßt.

20

In einer noch weiteren Ausführungsform kann vorgesehen sein, daß die Wechselwirkung vermittelnde Sequenz eine Konsensuspeptidsequenz umfaßt, welche ausgewählt ist aus der Gruppe, die die Sequenzen SEQ ID NO.: 8, 9, 10 11 und 12 umfaßt. In einer besonders bevorzugten Ausführungsform ist dabei vorgesehen, daß die Wechselwirkung vermittelnde Sequenz eine Konsensuspeptidsequenz gemäß SEQ ID NO.: 8 umfaßt.

In einer anderen Ausführungsform ist vorgesehen, daß die Wechselwirkung vermittelnde Sequenz am C-terminalen Ende der die Nukleinsäuresyntheseaktivität tragenden Sequenz liegt.



In einer weiteren Ausführungsform ist vorgesehen, daß zwischen der Wechselwirkung vermittelnden Sequenz und der die Nukleinsäureaktivität tragenden Sequenz ein Linker angeordnet ist.

Weiterhin kann in einer Ausführungsform der erfindungsgemäßen chimären Proteine das rekombinante chimäre Protein thermostabil sein.

10 In einer Ausführungsform der erfindungsgemäßen chimären Proteine kann vorgesehen sein, daß das Protein eine DNA Polymerase-Aktivität aufweist. Dabei ist besonders bevorzugt, wenn die Proteine eine 3'-5'- Exonukleaseaktivität aufweisen.

In einer weiteren Ausführungsform der erfindungsgemäßen chimären Proteine ist vorgesehen, daß das Protein eine RNA-Polymerase-Aktivität aufweist.

In einer weiteren Ausführungsform der erfindungsgemäßen chimären Proteine ist vorgesehen, daß das Protein eine reverse Transkriptase-Aktivität aufweist.

20 In einer noch weiteren bevorzugten Ausführungsform der erfindungsgemäßen chimären Proteine ist vorgesehen, daß sich die Einbaurate von dNTPs und ddNTPs durch die Nukleinsäuresyntheseaktivität um einen Faktor von weniger als 5 unterscheidet.

In einem weiteren Aspekt der Erfindung wird die Aufgabe gelöst durch einen Komplex umfassend

- a) ein erfindungsgemäßes rekombinantes chimäres Protein und
- b) ein Gleitklammerprotein.

30 In einer bevorzugten Ausführungsform des erfindungsgemäßen Komplexes ist vorgesehen, daß das Gleitklammerprotein eine Konsensuspeptidsequenz umfaßt, welche ausgewählt ist aus der Gruppe, die die Sequenzen gemäß SEQ ID NO.: 4, 5, 6 und 7 umfaßt.



In einer weiteren Ausführungsform ist vorgesehen, daß der Komplex weiter eine Nukleinsäure umfaßt.

In einem weiteren Aspekt wird die Aufgabe gelöst durch eine Nukleinsäure, die für ein erfindungsgemäßes chimäres Protein, insbesondere rekombinantes Protein codiert.

10 In einem noch weiteren Aspekt wird die Aufgabe gelöst durch eine die erfindungsgemäße Nukleinsäure umfassenden Vektor. In einer bevorzugten Ausführungsform handelt es sich bei dem Vektor um einen Expressionsvektor.

In einem weiteren Aspekt wird die Aufgabe gelöst durch eine Zelle, die den erfindungsgemäßen Vektor umfaßt.

Die Aufgabe wird weiter gelöst durch die Verwendung der erfindungsgemäßen chimären Protein zur Elongation von Nukleinsäuren.

20 Die Aufgabe wird weiter gelöst durch die Verwendung der erfindungsgemäßen chimären Protein zur Amplifikation von Nukleinsäuren.

Die Aufgabe wird weiter gelöst durch die Verwendung der erfindungsgemäßen chimären Protein zur reversen Transkription von RNA in DNA.

Die Aufgabe wird weiter gelöst durch die Verwendung der erfindungsgemäßen chimären Protein zur Sequenzierung von Nukleinsäuren, insbesondere von DNA.

30 In einem noch weiteren Aspekt wird die Aufgabe gelöst durch einen Kit, insbesondere Reagenzien-Kit zur Elongation und/oder Amplifikation und/oder reversen Tran-

skription und/oder Sequenzierung und/oder Markierung von Nukleinsäuren, der in einem oder mehreren getrennten Behältern umfaßt:

- a) ein erfindungsgemäßes, bevorzugt rekombinantes chimäres und/oder
- b) einen erfindungsgemäßen Komplex und
- c) bevorzugterweise gegebenenfalls zumindest einen Primer, Puffer, Nukleotide, Kofaktoren und/oder Pyrophosphatase.

10 In einer bevorzugten Ausführungsform des Kits ist vorgesehen, daß er zur Amplifikation von Nukleinsäuren neben den Substanzen a) und/oder b), Deoxynukleotide oder/und Derivate davon umfaßt.

In einer weiteren Ausführungsform umfaßt der Kit eine DNA Polymerase mit 3'-5' Exonukleaseaktivität umfaßt.

In einer noch weiteren Ausführungsform enthält der Kit zur Reversen Transkription Substanzen nach a) und/oder b), welche eine Reverse Transkriptase-Aktivität aufweisen, und bevorzugterweise Deoxynukleotide und/oder Derivate davon.

20 Bei dem erfindungsgemäßen Kit kann vorgesehen sein, daß er zur Sequenzierung neben Deoxynukleotiden oder Ribonukleotiden oder/und deren Derivaten, Dideoxynukleotide oder/und deren Derivate enthält.

In einem weiteren Aspekt wird die Aufgabe gelöst durch ein Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, wobei die zu elongierende Nukleinsäure oder zumindest ein Strang davon mit mindestens einem Primer unter Hybridisierungsbedingungen versehen wird, wobei der Primer hinreichend komplementär zu einem Teil der oder einem flankierenden Bereich der zu elongierenden Nukleinsäure ist und durch eine Polymerase in Anwesenheit von Nukleotiden eine Primer-Elongation erfolgt, wobei vorgesehen ist, daß als  
30 Polymerase ein erfindungsgemäßes chimäres, bevorzugt chimäres Protein

verwendet wird und daß bevorzugterweise ein Gleitklammerprotein in der Reaktion vorliegt.

In einem noch weiteren Aspekt wird die Aufgabe gelöst durch ein Verfahren zur Amplifikation einer Nukleinsäure, wobei vorgesehen ist, daß die zu amplifizierende Nukleinsäure mit mindestens zwei Primern unter Hybridisierungsbedingungen vesetzt wird, wobei ein jeder der beiden Primer jeweils komplementär zu einem Teil der oder einem flankierenden Bereich der zu amplifizierenden Nukleinsäure ist und  
10 durch eine Polymerase in Anwesenheit von Nukleotiden eine Primer-Elongation erfolgt, wobei als Polymerase ein erfindungsgemäßes chimäres Protein, insbesondere ein rekombinantes, verwendet wird und bevorzugterweise ein Gleitklammerprotein der Reaktion zugesetzt wird.

In einer bevorzugten Ausführungsform ist dabei vorgesehen, daß eine Polymerase-Ketten-Reaktion durchgeführt wird.

In einer ganz besonders bevorzugten Ausführungsform ist dabei vorgesehen, daß der Reaktionsansatz zwei DNA Polymerasen umfaßt, von denen mindestens eine eine  
20 3'-5' Exonukleaseaktivität aufweist, wobei die 3'-5'-Exonukleaseaktivität entweder durch das chimäre Protein oder durch eine weitere Polymerase dem Reaktionsansatz zugesetzt wird.

In einer noch weiteren Ausführungsform ist vorgesehen, daß zwei chimäre Proteine, insbesondere rekombinante, , wobei eines der Proteine ein solches ist, das eine DNA-Polymerase-Aktivität aufweist, und das andere ein solches ist, das eine 3'-5'-Exonukleaseaktivität aufweist.

Bei dem erfindungsgemäßen Verfahren zur Template-abhängigen Elongation kann in  
30 einer Ausführungsform vorgesehen sein, daß man zur Sequenzierung von Nukleinsäuren ausgehend von einem Primer, der zu einem der zu sequenzierenden Nukleinsäure benachbarten Bereich komplementär ist, eine Template-abhängige

Elongation bzw. Reverse-Transkription unter Verwendung von Deoxynukleotiden und Dideoxynukleotiden oder deren jeweiligen Derivaten gemäß der Methode von Sanger durchführt.

In einer weiteren Ausführungsform des erfindungsgemäßen Verfahren zur Template-abhängigen Elongation kann vorgesehen sein, daß bei der Elongation der Nukleinsäuren mindestens eine Markierungen eingefügt wird.

10 In einer besonders bevorzugten Ausführungsform ist dabei vorgesehen, daß ein Agens verwendet wird, das ausgewählt ist aus der Gruppe, die markierte Primer, markierte Deoxynukleotide und Derivate davon, markierte Dideoxynukleotide und Derivate davon und markierte Ribonukleotide und Derivate davon umfaßt.

In einem noch weiteren Aspekt wird die Aufgabe gelöst durch ein Verfahren zur Markierung von Nukleinsäuren durch Erzeugung einzelner Brüche in Phosphodiesterbindungen der Nukleinsäurekette und Ersatz eines Nukleotids an den Bruchstellen durch ein markiertes Nukleotid mit Hilfe einer Polymerase, wobei vorgesehen ist, daß als Polymerase ein erfindungsgemäßes chimäres Protein, insbesondere rekombinantes chimäres Protein eingesetzt wird.

20

In einem noch weiteren Aspekt wird die Aufgabe gelöst durch ein Verfahren zur Herstellung eines chimären Proteins, das eine Basissequenz und eine heterologe Wechselwirkung-bedingende Sequenz umfaßt und infolge der Wechselwirkung-bedingenden Sequenz mit einem Wechselwirkungspartner eine Bindung eingeht oder eine solche Bindung verstärkt wird, wobei

a) aus einem Wechselwirkungssystem umfassend ein als Donorprotein und ein als Akzeptorprotein bezeichnete Protein diejenige Sequenz des Donorproteins oder  
30 Akzeptorproteins ermittelt wird, die die Wechselwirkung zwischen beiden Wechselwirkungspartnern bedingt; und

b) die Wechselwirkung-bedingende Sequenz in ein vom Donorprotein und Akzeptorprotein verschiedenes aufnehmendes Protein, das die Basis sequenz umfaßt, eingebracht wird.

In einer Ausführungsform ist vorgesehen, daß das Donorprotein und das Akzeptorprotein einen Komplex bilden, der Nukleinsäure bindet.

10 In einer weiteren Ausführungsform ist vorgesehen, daß das Donorprotein und das Akzeptorprotein einen Komplex bilden, der eine Aktivität aufweist, die ausgewählt ist aus der Gruppe, die Polymeraseaktivität, DNA-Bindungsaktivität, RNA-Bindungsaktivität, 5'-3'-Exonukleaseaktivität, 3'-5'-Exonukleaseaktivität und Ligaseaktivität umfaßt.

In einer bevorzugten Ausführungsform ist vorgesehen, daß das Donorprotein ausgewählt ist aus der Gruppe, die Elongationsprotein, Gleitklammerproteine, Gleitklammerladerproteins und Kopplungsproteine umfaßt.

20 In einer weiteren Ausführungsform ist vorgesehen, daß das Akzeptorprotein ausgewählt ist aus der Gruppe, die Elongationsprotein, Gleitklammerproteine, Gleitklammerladerproteins und Kopplungsproteine umfaßt.

In einer noch weiteren Ausführungsform ist vorgesehen, daß das aufnehmende Protein ausgewählt ist aus der Gruppe, die Elongationsprotein, Gleitklammerproteine, Gleitklammerladerproteins und Kopplungsproteine umfaßt.

30 In einer Ausführungsform des erfindungsgemäßen Verfahrens kann vorgesehen sein, daß Schritt a) mehrfach wiederholt wird und aus den solchermaßen ermittelten Wechselwirkung-bedingenden Sequenzen eine Konsensussequenz ermittelt wird, die eine Wechselwirkung-bedingende Sequenz darstellt, und im Schritt b) als Wechselwirkung-bedingende Sequenz in ein vom Donorprotein und Akzeptorprotein

verschiedenes aufnehmendes Protein, das eine Basissequenz umfaßt, eingebracht wird.

In einem weiteren Aspekt wird die Aufgabe gelöst durch ein chimäres Protein, das nach dem erfindungsgemäßen Verfahren erhältlich ist. In einer bevorzugten Ausführungsform ist die Basissequenz ein Teil der Aminosäuresequenz eines Proteins, das ausgewählt ist aus der Gruppe, die Elongationsproteine, Gleitklammerproteine, Gleitklammerladerproteine und Kopplungsproteine umfaßt.

- 10 Schließlich wird die Aufgabe auch gelöst durch einen in vitro-Komplex zur Template-abhängigen Elongation von Nukleinsäure umfassend ein Gleitklammerprotein und ein Elongationsprotein, wobei mindestens eines der Proteine ein erfindungsgemäßes chimäres Protein ist. Besonders bevorzugt ist dabei eine Ausführungsform, bei der der Komplex thermostabil ist.

Im Zusammenhang mit der vorliegenden Erfindung gilt, daß ein jedes der erfindungsgemäßen chimären Proteine ein rekombinantes chimäres Protein sein kann.

20

Erfindungsgemäß wird die Aufgabe auch gelöst durch rekombinantes Chimäres Protein umfassend a) eine erste Domäne mit Nukleinsäuresyntheseaktivität, und b) eine Wechselwirkung vermittelnde Sequenz, dadurch gekennzeichnet, daß durch die Wechselwirkung vermittelnde Sequenz ein Komplex aus Nukleinsäuresyntheseaktivität und Gleitklammerprotein gebildet wird, wobei der Komplex verschieden ist von dem Komplex, den die Nukleinsäuresyntheseaktivität und/oder das Gleitklammerprotein mit ihren natürlichen Wechselwirkungspartner (n) ausbildet.

- 30 Es ist somit mittels dem erfindungsgemäßen rekombinanten chimären Protein möglich eine Nukleinsäuresyntheseaktivität an ein Gleitklammerprotein zu binden welches jene Nukleinsäuresyntheseaktivität, wie es beispielsweise in der Natur vor-



kommt nicht binden kann. Der natürliche Wechselwirkungspartner ist also ein Partner wie er innerhalb eines Organismus unter normalen physiologischen Bedingungen an der der Nukleinsäuresyntheseaktivität gebunden vorliegen kann.

Erfindungsgemäß wird die Aufgabe gelöst durch ein Chimäres Protein umfassend einen Nukleinsäuresyntheseaktivität aufweisenden Anteil, wobei der Anteil aus einem einen Nukleinsäuresyntheseaktivität aufweisenden Anteil und einen Wechselwirkung vermittelnden Anteil umfassenden Basisprotein stammt, und mindestens einen eine Wechselwirkung vermittelnden Anteil, wobei der Wechselwirkung vermittelnde Anteil verschieden ist von dem Wechselwirkung vermittelnden Anteil des Basisproteins, sowie ein Chimäres Protein umfassend einen Nukleinsäuresyntheseaktivität aufweisenden Anteil, wobei der Anteil aus einem Basisprotein stammt, das einen Nukleinsäuresyntheseaktivität aufweisenden Anteil aber keinen Wechselwirkung vermittelnden Anteil umfaßt, und einen Wechselwirkung vermittelnden Anteil.

Erfindungsgemäß wird die Aufgabe weiterhin gelöst durch ein Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, wobei die zu elongierende Nukleinsäure oder zumindest ein Strang davon mit mindestens einem Primer unter Hybridisierungsbedingungen versehen wird, wobei der Primer hinreichend komplementär zu einem Teil der oder einem flankierenden Bereich der zu elongierenden Nukleinsäure ist und durch eine Polymerase in Anwesenheit von Nukleotiden eine Primer-Elongation erfolgt, dadurch gekennzeichnet, daß als Polymerase ein erfindungsgemäßes rekombinantes chimäres Protein verwendet wird und daß bevorzugterweise ein Gleitklammerprotein in der Reaktion vorliegt.

Im Sinne der vorliegenden Erfindung ist unter „rekombinant“ zu verstehen, wenn das chimäre Protein beispielsweise gentechnisch hergestellt wird (siehe „Gentechnologie“, Römpp Basislexikon Chemie, Georg Thieme Verlag 1998) oder aber beispielsweise wenn es chemisch synthetisiert wird.

30

Weitere Ausführungsformen ergeben sich aus den Unteransprüchen.

Der Erfindung liegt die überraschende Erkenntnis zugrunde, daß es möglich ist ausgehend von einem eine Nukleinsäuresyntheseaktivität aufweisenden Protein, hierin im folgenden als Basisprotein bezeichnet, die Prozessivität dieses Proteins, genauer der Nukleinsäuresyntheseaktivität dadurch zu erhöhen, daß es mit einem die Prozessivität erhöhenden Faktor, wie einem Gleitklammerprotein, in Wechselwirkung tritt, mit dem das Basisprotein als solches nicht in Wechselwirkung treten kann, bzw. für den Fall, daß es in Wechselwirkung treten kann, dies nicht mit einer Erhöhung der Prozessivität verbunden ist. Dies wird dadurch bewerkstelligt, daß das die Nukleinsäuresyntheseaktivität-tragende Protein mit einer Gruppe von mehreren Aminosäuren versehen wird, typischerweise in der Form einer konsekutiven Aminosäuresequenz, die hierin als Wechselwirkung vermittelnde oder Wechselwirkung bedingende Sequenz bezeichnet wird, und damit eine Wechselwirkung zwischen dem besagten Protein und einem die Prozessivität erhöhenden Faktor erst möglich wird. Damit wird das oben geschilderte Problem des Standes der Technik bestehend in einer fehlenden Kompatibilität der eigentlich erwünschten Einzelkomponenten eines eine Nukleinsäuresyntheseaktivität tragenden Komplexes überwunden.

Ohne darauf beschränkt sein zu wollen, ergeben sich somit zumindest die folgenden drei Möglichkeiten, was die Ausbildung eine derartigen chimären, eine Nukleinsäuresyntheseaktivität tragenden Proteins anbelangt.

Das dem chimären Protein zugrundeliegende Basisprotein kann dabei in zwei grundsätzlichen Formen vorliegen. In der ersten Form umfaßt das Basisprotein lediglich die Nukleinsäuresyntheseaktivität, jedoch keine Sequenz (hierin auch als Domäne bezeichnet, die, insbesondere *in vivo*, eine Wechselwirkung zwischen der Nukleinsäuresyntheseaktivität und einem Prozessivitätsfaktor vermitteln kann oder könnte, insbesondere nicht dergestalt, als daß durch die Wechselwirkung eine Erhöhung der Prozessivität erreicht wird.

In einer zweiten Form umfaßt das Basisprotein eine Nukleinsäuresyntheseaktivität und zusätzlich eine Sequenz, die, insbesondere *in vivo*, eine Wechselwirkung zwischen der Nukleinsäuresyntheseaktivität und einem Prozessivitätsfaktor vermitteln kann oder könnte, insbesondere dergestalt, als daß durch die Wechselwirkung eine Erhöhung der Prozessivität erreicht wird.

Das erfindungsgemäße chimäre Protein kann ausgehend von den verschiedenen Formen des Basisproteins in verschiedenen grundsätzlichen Ausführungsformen vorliegen.

10

(i) Eine erste Ausführungsform des chimären Proteins sieht vor, daß die erste Form des Basisproteins verwendet wird und daran eine eine Wechselwirkung mit einem die Prozessivität der Nukleinsäuresyntheseaktivität erhöhenden Faktor vermittelnde Sequenz angefügt wird. Diese Anfügung erfolgt typischerweise dadurch, daß die Wechselwirkung vermittelnde Sequenz sich der Sequenz der Nukleinsäuresyntheseaktivität, gegebenenfalls getrennt durch einen Linker, anschließt. In dieser ersten Ausführungsform des chimären Proteins wird somit das Basisprotein um eine eine Wechselwirkung vermittelnde Sequenz ergänzt.

20

(ii) In einer zweiten Ausführungsform des chimären Proteins wird die zweite Form des Basisproteins verwendet. Dabei wird die dem Basisprotein immanente eine Wechselwirkung mit einem die Prozessivität der Nukleinsäuresyntheseaktivität erhöhenden Faktor vermittelnde Sequenz durch eine andere eine Wechselwirkung mit einem die Prozessivität der Nukleinsäuresyntheseaktivität erhöhenden Faktor vermittelnde Sequenz ausgetauscht. Diese andere Sequenz kann dabei von einem anderen Gen des gleichen Organismus, vom gleichen Gen eines anderen Organismus oder von einem anderen Gen eines anderen Organismus stammen und ist somit in einem jeden Fall eines anderen Ursprungs. Infolgedessen wird durch eine derartige Konstruktion, ggf. erstmals, ggf. aber auch nur in verstärktem Maße, eine Wechselwirkung zwischen der Nukleinsäuresyntheseaktivität des Basisproteins und einem die Prozessivität der Nukleinsäuresyntheseaktivität erhöhenden Faktor möglich.

30

(iii) In einer dritten Ausführungsform des chimären Proteins wird die zweite Form des Basisproteins verwendet. Dabei wird die dem Basisprotein immanente eine Wechselwirkung mit einem die Prozessivität der Nukleinsäuresyntheseaktivität erhöhenden Faktor vermittelnde Sequenz ergänzt durch eine andere eine Wechselwirkung mit einem die Prozessivität der Nukleinsäuresyntheseaktivität erhöhenden Faktor vermittelnde Sequenz. Diese andere Sequenz kann dabei von einem anderen Gen des gleichen Organismus, vom gleichen Gen eines anderen Organismus oder von einem anderen Gen eines anderen Organismus stammen und ist somit in einem  
10 jeden Fall eines anderen Ursprungs. Infolgedessen wird durch eine derartige Konstruktion, ggf. erstmals, ggf. aber auch nur in verstärktem Maße, eine Wechselwirkung zwischen der Nukleinsäuresyntheseaktivität des Basisproteins und einem die Prozessivität der Nukleinsäuresyntheseaktivität erhöhenden Faktor möglich.

Für alle der oben geschilderten Ausführungsformen des chimären Proteins kann man feststellen, daß sich typischerweise die (Aminosäure-)Sequenz eines derartigen chimären Proteins von der Sequenz des Basisproteins unterscheidet. Eine weitere, allerdings nicht obligatorische Folge kann darin gesehen werden, daß der durch das chimäre Protein ausgebildete Komplex aus Nukleinsäuresyntheseaktivität aufwei-  
20 sendem Protein und Prozessivität erhöhendem Faktor verschieden ist von dem Komplex aus Basisprotein und Prozessivität erhöhendem Faktor.

Als Basisproteine eignen sich alle in der Einleitung hierin erwähnten Proteine, Polymerasen und Elongationsproteine, deren Offenbarung hiermit durch Bezugnahme hierin aufgenommen werden. Gleiches gilt für die anderen Komponenten, insbesondere derjenigen des Replikationsapparates, wie Prozessivitätsfaktoren, Gleitklammerproteine und Wechselwirkung-vermittelnde oder Wechselwirkung-bedingende Sequenzen. Schließlich gelten auch für diesen Teil der Offenbarung die in der Einleitung gegebenen Definitionen.

Wenngleich die Nukleinsäuresyntheseaktivität aus einer Vielzahl von Organismen stammen kann so ist es bevorzugt, wenn sie z.B. aus dem Organismus *Carboxydotherrnus hydrogenofornans* (European Patent Application EP 0 834 569 A1) oder einem der Organismen wie z.B. *Therrnus aquaticus*, *Therrnus caldophilus*, *Therrnus chliarophilus*, *Therrnus filiformis*, *Therrnus flavus*, *Therrnus oshimai*, *Therrnus ruber*, *Therrnus scotoductus*, *Therrnus silvanus*, *Therrnus species ZO5*, *Therrnus species sp. 17*, *Therrnus therrnusphilus*, *Therotoga maritima*, *Therotoga neapolitana*, *Therrmosipho africanus*, *Anaerocellum therrmophilum*, *Bacillus caldotenax*, oder *Bacillus stearothermophilus* stammt.

10

Sofern hierin von Sequenz gesprochen wird, wird damit in der Regel Bezug genommen auf eine Aminosäuresequenz. Nukleinsäuresequenzen werden in der Regel direkt als Nukleinsäuresequenzen angesprochen.

Die verschiedenen für die erfindungsgemäßen rekombinanten chimären Proteine codierenden Nukleinsäuren können von den Fachleuten auf dem Gebiet mittels des genetischen Codes leicht ermittelt und sodann synthetisiert werden. Den Fachleuten sind ebenfalls geeignete Vektoren zur Klonierung und Expression der erfindungsgemäßen rekombinanten chimären Proteine sowie Verfahren zu deren bevorzugterweise rekombinanten Produktion bekannt (siehe z.B. Maniatis et al.; aaO).

20

Den Fachleuten ebenfalls bekannt sind Verfahren zur Herstellung von Antikörpern, einschließlich monoklonalen Antikörpern, die gegen die erfindungsgemäßen rekombinanten chimären Proteine gerichtet sind.

Das erfindungsgemäße rekombinante chimäre Protein kann zur Elongation von Nukleinsäuren verwendet werden, z. B. zur Polymerase-Ketten-Reaktion, DNA-Sequenzierung, zur Markierung von Nukleinsäuren und anderen Reaktionen, die die in vitro Synthese von Nukleinsäuren beinhalten.

30



Ein weiterer Gegenstand der vorliegenden Erfindung ist daher ein Verfahren zur Template-abhängigen Elongation, wobei die zu elongierende Nukleinsäure oder zumindest ein Strang davon mit mindestens einem Primer unter Hybridisierungsbedingungen versehen wird, wobei der Primer hinreichend komplementär zu einem Teil der oder einem flankierenden Bereich der zu elongierenden Nukleinsäure ist und durch eine Polymerase in Anwesenheit von Nukleotiden eine Primer-Elongation erfolgt, wobei als Polymerase ein erfindungsgemäßes rekombinantes chimäres Protein verwendet wird und in einer bevorzugten Ausführungsform weiterhin ein Gleitklammerprotein in der Reaktion bzw. im Reaktionsansatz vorliegt.

Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, bei denen die Elongation ausgehend von einem Primer erfolgt, der an die Template-Nukleinsäure anhybridisiert wurde und ein freies 3'-OH-Ende für die Elongation zur Verfügung stellt, sind dem Fachmann bekannt. Zur Amplifikation wird insbesondere eine Polymerasekettenreaktion durchgeführt. Hierbei wird üblicherweise von einer doppelsträngigen DNA-Sequenz ausgegangen, von welcher ein bestimmter Zielbereich amplifiziert werden soll. Hierbei werden zwei Primer eingesetzt, welche zu der Zielsequenz flankierenden Bereichen auf jeweils einem Teilstrang des DNA-Doppelstranges komplementär sind. Zur Anhybridisierung der Primer werden allerdings die DNA-Doppelstränge zuerst denaturiert, insbesondere thermisch aufgeschmolzen. Nach Anhybridisierung der Primer erfolgt eine Elongation mittels der Polymerase, daraufhin wird nochmals denaturiert und damit die neu gebildeten DNA-Stränge von den Template-Strängen getrennt, woraufhin für einen weiteren Elongationszyklus neben den ursprünglichen Templatesträngen auch die im ersten Schritt gebildeten Nukleinsäurestränge als Template zur Verfügung stehen, diese jeweils erneut mit Primern hybridisiert werden und eine erneute Elongation stattfindet. Diese Vorgehensweise wird zyklisch durchgeführt unter jeweils thermischer Denaturierung als Zwischenschritte.



Auch möglich ist die Anwendung des erfindungsgemäßen rekombinanten chimäre Proteins in der reversen Transkription, wobei entweder das erfindungsgemäße Protein selbst Reverse-Transkriptase-Aktivität besitzt, oder aber zusätzlich ein geeignetes Enzym hinzugefügt wird, das eine Reverse-Transkriptase-Aktivität aufweist, unabhängig davon ob der thermostabile in vitro Komplex selbst eine Reverse-Transkriptase-Aktivität hat.

10 Zur erfindungsgemäß bevorzugten reversen Transkription von RNA in DNA wird ebenfalls ein erfindungsgemäßes rekombinantes chimäres Protein eingesetzt, wobei dessen Nukleinsäuresyntheseaktivität selbst eine Reverse Transkriptase-Aktivität aufweist. Diese Reverse Transkriptase-Aktivität kann die einzige Polymeraseaktivität sein, kann aber auch zusätzlich zu einer vorhandenen 5'-3'-DNA-Polymeraseaktivität vorliegen. Eine erfindungsgemäß bevorzugte Ausführungsform des rekombinanten chimären Proteins umfaßt das Elongationsprotein aus dem Organismus *Carboxydotherrhus hydrogeniformans* wie offenbart in EP-A 0 834 569 stammt.

20 Eine weitere bevorzugte Verwendung des erfindungsgemäßen rekombinanten chimären Proteins ist die Sequenzierung von Nukleinsäuren ausgehend von mindestens einem Primer, der zu einem Teil der zu sequenzierenden Nukleinsäure hinreichend komplementär ist, wobei wiederum eine Template-abhängige Elongation oder aber bei Sequenzierung von RNA eine Reverse Transkription unter Verwendung von Deoxynukleotiden und Dideoxynukleotiden gemäß der Methode von Sanger durchgeführt wird. Als Deoxynukleotide oder Dideoxynukleotide werden im Rahmen dieser bevorzugten Ausführungsform auch die oben beschriebenen jeweiligen Derivate als geeignet angesehen. Insbesondere ist es für die erfindungsgemäßen Verfahren zur Elongation von Nukleinsäuren bevorzugt, daß die gebildeten Nukleinsäuren markiert werden. Hierzu ist es möglich, markierte Primer und/oder markierte Deoxynukleotide oder/und markierte Dideoxynukleotide und/oder markierte Ribonukleotide oder jeweils entsprechende Derivate, wie sie oben bereits beispielsweise beschrieben sind,  
30 einzusetzen.

Ein weiterer Gegenstand der vorliegenden Erfindung ist ein Verfahren zur Markierung von Nukleinsäuren durch Einfügung einzelner Brüche in Phosphodiesterbindungen der Nukleinsäurekette und Ersatz eines Nukleotids an den Bruchstellen durch ein markiertes Nukleotid mit Hilfe einer Polymerase, wobei als Polymerase ein erfindungsgemäßer thermostabiler in vitro-Komplex eingesetzt wird.

Ein solches Verfahren, allgemein Nick-Translation genannt, ermöglicht eine einfache Markierung von Nukleinsäuren. Alle oben bereits beschriebenen markierten Ribonukleotide oder Deoxyribonukleotide oder Derivate davon sind hierfür geeignet, 10 solange die Polymerase sie als Substrat akzeptiert.

Die Erfindung wird anhand der folgenden Beispiele und Figuren weiter erläutert, aus denen sich weitere Vorteile und Ausführungsformen ergeben, dabei zeigt

- Fig. 1 vier Sequenzalignments von verschiedenen Elongationsporteindomänen;
- Fig. 2 eine schematische Darstellung des erfindungsgemäßen rekombinanten chimären Proteins;
- 20 Fig. 3 vier Konsensussequenzen von Gleitklammerproteinen;
- Fig. 4 in tabellarischer Form das Ergebnis eines Hefe-Zweihybrid-system;
- Fig. 5, 5 A-C Alignments mehrerer konservierter Regionen der Wechselwirkung bedingenden Sequenz;
- Fig. 6 A die gesamte Sequenz einer Ausführungsform eines erfindungsgemäßen rekombinanten chimären Proteins;
- Fig. 6 B eine graphische Darstellung des prinzipiellen Aufbaus eines erfindungsgemäßen chimären Proteins;
- Fig. 7 das Ergebnis einer unter Verwendung eines erfindungsgemäßen chimären Proteins durchgeführten Polymerase-Kettenreaktion;
- 30 Fig. 8 A eine Darstellung aller Fragmente von Af0497, die mit dem Gleitklammerprotein von Archaeoglobus fulgidus (Af0335) wechselwirken;

- Fig. 8 B ein Alignment C-terminaler Sequenzen verschiedener Gene von *Archaeoglobus fulgidus*;
- Fig. 9 das Ergebnis eines Hefe-Zweihybridexperimentes, das die Wechselwirkung zwischen einem Elongationsprotein und einem Gleitklammerprotein darstellt; und
- Fig. 10 das Ergebnis der Amplifikation genomischer DNA unter Verwendung eines erfindungsgemäßen rekombinanten chimären Proteins.

10 **Beispiel 1: Anwendung des Hefe-Zweihybridsystems zur Bestimmung der für die Wechselwirkung zwischen Replikationsfaktoren und Gleitklammerprotein bedeutsamen Aminosäuren**

Zur Bestimmung der Binderegion von Gleitklammerproteinen (hierin im folgenden als GKP bezeichnet) und Replikationsfaktoren (hierin im folgenden als RF bezeichnet) wird das Hefe-Zweihybridsystem (Fields S., Song O., Nature 1989 Jul 20; 340 (6230):245-6) verwendet. Die Gene für GKPs und RFs werden in den Vektoren pGBT9 und pGAD424 (CLONTECH Laboratories, Inc.) exprimiert, so daß Fusionsproteine mit der DNA-Bindedomäne bzw. der Aktivierungsdomäne von GAL4 entstehen: DNA wird mittels der bekannten Verfahren aus dem Organismus

20 *Archaeoglobus fulgidus* (DSM No. 4304) gereinigt. Die Aufzucht der Mikroorganismen geschah hierbei durch die DSM (Deutsche Sammlung für Mikroorganismen). Dann werden die offenen Leseraster der Gene mittels PCR aus genomischer DNA von *Archaeoglobus fulgidus* amplifiziert. Die solchermaßen erhaltene DNA wird in die Vektoren pGBT9 und pGAD424 kloniert. Andere Vektoren, die im Zweihybridsystem verwendet werden, eignen sich auch für das im Folgenden beschriebene Verfahren (darunter sind zum Beispiel pAD-GAL4-2.1, pBD-GAL4, pBD-GAL4 Cam, pCMV-AD, pCMV-BD, pMyr, pSos, pACT2, pAS2-1, pHISi, pLexA, pM, pHISi-1, pB42AD, pVP16, pGAD10, pGBKT7, pLacZi, p8op-lacZ, pGAD GH, pGilda, pAD GL, pGADT7, pGBDU, pDBLeu, pPC86, pDBTrp, die erhältlich sind von CLONTECH Laboratories,

Inc.). Zusätzlich werden veränderte Klone in die gleichen Vektoren kloniert. Bei den veränderten Klonen handelt es sich um Deletionsmutationen und um Mutationen betreffend einzelne oder mehrere Aminosäuren des GKP und/oder des RF. Anschließend wird die Fähigkeit der veränderten Klone, im Zweihybridsystem miteinander zu interagieren, gemessen. Dadurch wird es möglich, Domänen oder sogar einzelne Aminosäurereste zu bestimmen, die für die Interaktion wichtig oder essentiell sind.

Es gibt eine Reihe von Methoden, Deletionen in ein Gen einzuführen. Eine Methode zur  
10 Herstellung von Deletionen besteht darin, DNA-Fragmente herzustellen, die die Gene für das GKP oder den RF enthalten. Diese Fragmente werden entweder durch PCR oder durch Restriktionsverdau aus einem geeigneten Vektor gewonnen. Des weiteren wird genomische DNA des Organismus verwendet, aus dem die beiden Proteine stammen. Diese verschiedenen DNA-Fragmente und die genomische DNA werden durch Ultraschallbehandlung zerkleinert und Fragmente, die Längen zwischen der Gesamtlänge der Gene und etwa 100 Basen aufweisen, werden durch präparative Agarosegelelektrophorese gereinigt. Die Enden der Fragmente werden  
20 durch Behandlung mit einem geeigneten Enzym (Klenow, Pwo-Polymerase, oder andere) aufgefüllt bzw. abverdaut, so daß beide Stränge die gleiche Länge haben und somit glattendig ("blunt") sind. Diese Fragmente werden nun durch Ligation in einen mit SmaI geschnittenen (oder anders linearisierten und blunt gemachten) pGAD424 und pGBT9 Vektor, oder andere geeignete Vektoren, eingebaut. So entstehen Banken von einer Vielzahl von Subfragmenten der Gene für GKP und RF, jeweils in beiden Vektoren des Zweihybridsystems. Die DNA dieser Banken wird nach Transformation in *Escherichia coli* vermehrt und gereinigt. Im nächsten Schritt werden die beiden Banken in einen für das Zweihybridsystem geeigneten haploiden Hefestamm transformiert, so daß die Fusionen aus GKP bzw. RF mit der GAL4 DNA-Bindedomäne in einem Stamm mit einem anderen Paarungstyp vorliegen als die Fusionen aus RF bzw. GKP der GAL4 Aktivierungsdomäne. Durch Paarung der  
30 beiden Stämme werden nun diploide Zellen erzeugt, in denen die Plasmide aus beiden Banken in derselben Zelle vorliegen. Als Stamm eignet sich PJ69-4 (James P,

Halladay J, Craig EA, Genetics 1996 Dec;144(4):1425-36) oder jeder andere Stamm mit einem für das Zweihybridsystem geeigneten Genotyp. Zellen, in denen die Reportergene aktiviert werden, werden isoliert, und die Plasmid-DNA daraus durch Präparation gewonnen, oder die Inserts durch PCR spezifisch vermehrt. Die Sequenzen der Fusionsfragmente werden durch DNA-Sequenzanalyse ermittelt. Die Bindungsregion (oder die Bindungsregionen) wird (werden) durch Bestimmung jener Bereiche ermittelt, die in allen Klonen (oder die in bestimmten Gruppen von Klonen immer gefunden werden), bei denen es zu einer Wechselwirkung der beiden RF und GKP umfassenden Fusionsproteine kommt. Diese Bestimmung wird für jedes der

10 Gene viermal durchgeführt: einmal mit der Bank im Vektor mit der Aktivierungsdomäne (präferentiell: pGAD424) und einmal mit der Bank im Vektor mit der DNA-Bindedomäne (präferentiell: pGTB9), jeweils einmal mit dem Gesamtlängenklon des Bindepartners und einmal mit der Genbank der Fragmente.

Durch die beschriebenen Versuche ist es möglich, minimale Binderegionen zu definieren, die man für die Konstruktion von rekombinanten chimären Proteinen mit Affinität für GKP nutzen kann. Zur genaueren Charakterisierung der Binderegion werden Aminosäuremutationen eingeführt und deren Wirkung auf die Bindeeigenschaft getestet. Dadurch wird der Bereich weiter eingeeengt und die an der

20 Bindung beteiligten Positionen im Protein bestimmt. In einem parallelen Ansatz wird durch das Einführen zufälliger Mutationen in die Binderegion und der Analyse der Wirkung auf die Bindung ebenfalls die an der Bindung beteiligten Positionen bestimmen. In beiden Fällen wird bestimmt, ob die Mutationen das Protein instabil machen und also einen indirekten Effekt haben, oder ob sie keine Auswirkungen auf die Stabilität des Proteins haben und so einen direkten Effekt haben.

Die hierfür nötigen experimentellen Verfahren sind beschrieben in Maniatis et al. (*Molecular Cloning (2<sup>nd</sup> edition, 3 Volume set): A laboratory Manual*, Cold Spring Harbor Laboratory Press, N.Y.(1989)), in Ausubel et al (*Current Protocols in Molecular Biology*, John Wiley and Sons (1988)), in Abelson, J.N., and Simon, M.I. (Editoren) 1991

30 (*Methods in Enzymology, Volume 194, Guide to Yeast Genetics and Molecular*



Biology, Academic Press, ) oder in Adams, A. Gottschling, D.E. Kaiser, CA, and Stearns, T., 1997, (Methods in Yeast Genetics, Cold Spring Harbour Laboratory Press). Die Handhabung des Zweihybridsystems (two-hybrid system) erfolgte gemäß der Anleitung der Firma Clontech (yeast protocols handbook, PT3024-1).

### **Beispiel 2: Weitere Verfahren zur Kartierung der Bindungsregion von GKP und RF**

10 Es wurden GKP und RF als rekombinante Proteine gewonnen. Jeweils eines der beiden Proteine wurde an einem Träger immobilisiert, dann wurde das jeweils andere Protein in ungebundener Form zugeben, um die Bindung an den immobilisierten Partner erlauben. Freies, d.h. nicht gebundenes Material wurde ausgewaschen, und das gebundene Protein z.B. durch Denaturierung eluiert. Die Menge des gebundenen Proteins ist ein Mass für die Bindung (z.B. angewandt in Anderson D, Koch CA, Grey L, Ellis C, Moran MF, Pawson T, Science 1990 Nov 16;250(4983):979-82).

20 Die Analyse von Deletionsmutanten der Proteine und von Mutanten mit veränderter Aminosäuresequenz erlaubte die Kartierung der Bindungsregion. Deletionen können durch proteolytischen Verdau oder mittels der rekombinanten Expression von deletierten Genen erzeugt werden. In gleicher Weise kann man durch Co-Immunpräzipitation aus Zellextrakten, die deletierte Formen der Proteine enthalten, die Binderegion bestimmen. Die Kompetition der Bindung durch Peptide kann Aufschluß über die Sequenz der Binderegion geben.

Eine weitere Möglichkeit zur Kartierung der Binderegion besteht darin, Peptide herzustellen und die Bindung der Peptide an das jeweils andere Protein zu messen. Die Sequenz der Peptide kann zufällig (Songyang Z, Prog Biophys Mol Biol 1999;71(3-4):359-72) oder auf der Aminosäuresequenz des Proteins beruhen, dessen Binderegion identifiziert werden soll (petide scans, Brix J, Rudiger S, Bukau B, Schneider-Mergener J, Pfanner N, J Biol Chem 1999 Jun 4;274(23):16522-30). Solche Peptide

30

können durch chemische Synthese oder durch andere Methoden wie phage display hergestellt und zur Bindung angeboten werden. Ein Beispiel ist in Fig. 3 gezeigt. Hier wurden durch Alignments verschiedene Konsensussequenzen ermittelt, die zur Herstellung derartiger Peptide geeignet sind.

Ein weiteres Verfahren besteht darin, Antikörper herzustellen gegen Epitope des Proteins, dessen Binderegion identifiziert werden soll. Wenn der Antikörper die Bindung inhibiert, so überschneidet sich das Epitop es Antikörpers mit der Binderegion (z.B. Fumagalli S, Totty NF, Hsuan JJ, Courtneidge SA, Nature 1994 Apr 10 28;368(6474):871-4).

Schließlich besteht eine Möglichkeit zur Bestimmung oder Kartierung der Binderegion (Wechselwirkung-bedingende Sequenz) darin, die Feinstruktur des Komplexes mit den Methoden der Röntgenstrukturanalyse, der Nuclear Magnetic Resonance oder der Elektronenmikroskopie aufzuklären.

**Beispiel 3: Studium der Wechselwirkung von Proteinen aus *Archaeoglobus fulgidus* unter Verwendung des Hefe-Zweihybridsystems (Y2H)**

20 Die kodierenden Bereiche von Genen aus *Archaeoglobus fulgidus*, deren Genprodukte in vitro als Wechselwirkung-bedingende Sequenz und/oder Nukleinsäuresynthseaktivität verwendet werden können, wurden mittels PCR amplifiziert, in die Vektoren pGBT9 (vertikale Spalten der Fig. 9) und pGAD424 (horizontale Reihen der Fig. 9) kloniert und durch gap-repair in Hefe PJ69-4a (für pGAD424) und PJ69-4alpha (für pGBT9) als Hybridproteine zur Expression gebracht. Ebenso wurde eine positive Kontrolle mittels PCR amplifiziert, in die Vektoren pGBT9 und pGAD424 (siehe auch horizontale Reihen der Fig. 9) kloniert und durch gap-repair in den Hefestamm PJ69-4a (für pGAD424) und PJ69-4alpha (für pGBT9) als Hybridproteine zur Expression gebracht. Diploide Zellen, die beide Vektoren enthalten, wurden  
30 durch Paarung erzeugt. Die Expression von drei unabhängigen Reportern (*HIS3*,

*ADE2* und *MEL1*) wurde gemessen. Die Expression des *HIS3* bzw. *ADE2* Gens führt zu einer Histidin- bzw. Adenin-Prototrophie der Zellen. Transkription des *MEL1* Gens führt zur Produktion des Beta-Galaktosidase Enzyms, dessen Aktivität sich z.B. über eine Farbreaktion nachweisen läßt. In dem in Fig. 9 gezeigten Experiment wachsen jene Zellen in einem Histidin- und Adenin-Mangelmedium, die beide Vektoren tragen und in denen zudem die Expressionsprodukte dieser beiden Vektoren aneinander binden. Die Interaktion der Fusionsproteine bewirkt die Rekonstitution eines funktionellen Transkriptionsfaktors, der die Transkription der Reportergene initiiert. Die Produktion der auf den Reportergenen codierten Proteine führt zu einer Aufhebung der Histidin- und Adenin-Auxotrophie. Dies wird dadurch bedingt, daß in Folge der Bindung der Expressionsprodukte eine Transkription initiiert wird, die dazu führt, daß die Zellen wachsen können.

Alle hier positiven Klone waren auch positiv bezüglich der Expression des *MEL1*-Gens. Die Handhabung des two-hybrid systems erfolgte gemäß der Anleitung der Firma Clontech (yeast protocols handbook, PT3024-1).

#### **Beispiel 4: Ermittlung der Wechselwirkung bedingenden Sequenz von Af0497**

Zur Ermittlung der Wechselwirkung-bedingenden Sequenz auf dem Protein Af0497 (ein Elongationsprotein), die die Wechselwirkung mit dem Gleitklammerprotein bedingt, wurde zunächst die DNA des Gens für Af0497 durch Restriktion eines geeigneten Vektors, welcher eine Elongationsprotein beinhaltet welches eine Wechselwirkung-bedingende Sequenz umfaßte gewonnen. Diese DNA wurde dann durch Ultraschall fragmentiert und in die beiden Vektoren pGAD424 und pGBDU ligiert. So entstanden Banken von verschiedenen Fragmenten des Gens in den beiden Vektoren. Als nächstes wurde unter Verwendung des Hefe-Zweihybridsystems ermittelt, welches dieser Fragmente eine Wechselwirkung mit Af0335 bedingen kann. Dazu wurden beide Banken in geeignete Hefestämme (pGAD424: PJ69-4a, PGBDU: PJ69-4alpha) transformiert. Durch Paarung der Hefestämme wurden dann diploide

Zellen erzeugt, die sowohl die Banken (Af 0497) als auch das Gleitklammerproteins in einem für die Verwendung der Zweihybridsystems passenden Vektor enthalten. Alle Klone, die den HIS2 Reporter (siehe oben) aktivierten, wurden auf Platten ohne Histidin isoliert, und die Inserts wurden durch PCR von der Hefekolonie selektiv amplifiziert. Die Sequenz der Inserts wurde ermittelt und ihre Position am Gen Af0497 bestimmt. Das Ergebnis ist in Figure 8A graphisch dargestellt: Alle gefunden Klone mit Ausnahme eines Klons beinhalteten das carboxyterminale Ende des Elongationsproteins, welches eine oder mehrere Wechselwirkung-bedingende Sequenz umfaßt. Der kleinste Klon bestand aus 51 Aminosäuren.

10

Um zu ermitteln, ob die Wechselwirkung-bedingende Sequenz des homologen Proteins eines anderen Organismus ebenfalls am carboxyterminalen Ende des Proteins beinhaltet ist, wurde das Experiment mit dem Polymeraseprotein und dem Gleitklammerprotein aus *Pyrococcus horikoshii* wiederholt. Dazu wurden Banken der großen Untereinheit der Polymerase aus *P. horikoshii* hergestellt und wie mittels des oben beschriebenen Verfahrens auf Interaktion mit dem Gleitklammerprotein aus *P. horikoshii* getestet. Wieder umfaßten die interagierenden Fragmente das carboxyterminale Ende des Elongationsproteins.

20

Um zu überprüfen, ob die Wechselwirkung zwischen dem carboxyterminalen Ende des Proteins Af0497 und dem Gleitklammerprotein spezifisch ist, wurde getestet, ob das carboxyterminal Proteinfragment von Af0497 mit anderen Proteinen aus *Archaeoglobus fulgidus* und anderen Organismen ebenso wechselwirkt. In keinem Fall konnte eine Wechselwirkung gemessen werden, was zeigt, daß die Bindung sehr spezifisch für das Gleitklammerprotein ist.

So bindet die Polymerase Taq nicht an das GKP (PCNA aus *Archaeoglobus fulgidus*), und es konnte im Hefe-Zweihybridsystem keinerlei Wechselwirkung gemessen werden. Um zu testen, ob das carboxyterminale Ende des Proteins Af0497 eine

30

Wechselwirkung mit einem anderen Protein, an das es fusioniert ist, bedingen kann, wurde ein Fusionsprotein aus Taq und den carboxyterminale 51 Aminosäuren von

Af0497 hergestellt. Dieses Protein wurde im Hefe- Zweihybridsystem auf Wechselwirkung getestet. Wie in Fig. 4 dargestellt, bewirkte die Verpflanzung der Wechselwirkung-bedingenden Sequenz von Af0497 an Taq eine spezifische Wechselwirkung von Taq mit dem Gleitklammerprotein Af0335. Die Ergebnisse der entsprechenden Y2H-Experimente sind in Fig. 4 dargestellt.

Das beweist, daß die Eigenschaft dieses Fragments (das carboxyterminale Ende des Elongationsproteins Af0497) eine Wechselwirkung mit dem Gleitklammerprotein Af0335 zu bedingen, auf ein anders Protein, insbesondere auf eine andere Polymerase, verpflanzt oder transferiert werden kann, um dort eine Wechselwirkung mit dem eigentlich für Af0497 spezifischen Gleitklammerprotein hervorzurufen.

Es wurde bei sechs Proteinen (siehe auch Fig. 5 und Fig. 5B) aus *A. fulgidus* eine Interaktion mit PCNA gemessen. Daraus ergibt sich, daß alle diese Proteine ähnlich Wechselwirkung-bedingende Sequenzen beinhalten, die die Wechselwirkung mit PCNA vermitteln. So z.B. die Polymerase delta grosse Untereinheit (Af 0497, TREMBL Nummer: O29753), die Polymerase delta kleine Untereinheit (Af 1790, TREMBL Nummer: O28484), DP2 (Af 1722, TREMBL Nummer: O28552), RPA2 (Replication factor A), RFC2 (Replication factor C) und PCNA (Af 0335, TREMBL Nummer: O29912)

Solche Wechselwirkung-bedingende Sequenzen sind in den letzten 50 Aminosäuren des Proteins Af0497 enthalten. Bei einer Untersuchung der Proteine, die mit PCNA interagieren, ergab sich, daß alle ein Motiv enthalten, das knapp vor dem carboxyterminalen Ende der Aminosäuresequenz angeordnet ist. Eine Aufstellung der verwandten Sequenzen ist in Fig. 8B dargestellt, wobei ein schwarzer Balken den konservierten Bereich kennzeichnet.

Dieses Motiv ist auch in anderen Organismen und Genen konserviert. Fig. 5 zeigt das Ergebnis der Suche nach solchen Wechselwirkung-bedingende Sequenzen sowie die daraus generierten Konsensussequenzen.



**Beispiel 5: Steigerung der Effizienz einer PCR bei Verwendung von PCNA als GKP und eines erfindungsgemäßen chimären Elongationsproteins**

Dieses Beispiel zeigt den Einfluß von PCNA auf die Effizienz einer PCR – Reaktion. In dieser PCR – Reaktion wurde ein 463 bp großes Fragment ausgehend von Plasmid-DNA amplifiziert. Als Polymerase wurde das Taqfusionsprotein, wie es auch in Fig. 6A dargestellt ist, verwendet. Die Reaktionsbedingungen für die PCR waren wie folgt: 0,4mM jedes dNTP (pH 8,3) und 20 pmol von jedem Primer in einer Reaktion.

10 Zur Anwendung kam ein erster Primer (SEQ ID NO.: 13) mit der Sequenz 5'-AGGGCGTGGTGCGGAGGGCGGT-3' und ein zweiter Primer (SEQ ID NO.: 14) mit der Sequenz 5'-TCGAGCGGCCCGCCCGGGCAGGT-3'.

Es zeigt sich, daß die Fusion einer 50 Aminosäuren großen Domäne an den C-Terminus der nativen Taq DNA Polymerase keinen nachteiligen Einfluß auf die Polymerase – Eigenschaften hat, da das Taqfusionsprotein per se schon ein Produkt in einer PCR – Reaktion liefert. Mittels dieser Domäne kann nun das Gleitklammerprotein, PCNA aus *Archaeoglobus fulgidus*, mit dem Taqfusionsprotein interagieren und in seiner Eigenschaft als Prozessivitätsfaktor die Effizienz der PCR – Reaktion erhöhen, was sich in einer deutlich höheren Ausbeute an PCR Produkt widerspiegelt. Das

20 Ergebnis ist auch in Fig. 7 dargestellt.

**Beispiel 6: Hefe-Zweihybridsystem zum Nachweis der Wechselwirkung zwischen dem Elongationsprotein Af 0497 von *Archaeoglobus fulgidus* und dem Gleitklammerprotein Af 0335 von *Archaeoglobus fulgidus***

Fig. 9 zeigt die Ergebnisse eines Y2H Experimentes, wobei die Reihe A mit Zellen bestückt ist, die den leeren pGAD424 Vektor (Clontech, Palo Alto, USA) tragen, so daß eine Transkriptions-Aktivierungsdomäne exprimiert wird; die Reihe B mit Zellen

bestückt ist, die den pGAD424 Vektor tragen, von dem das *Sacharomyces cerevisiae* Gen CDC48 als Fusionsprotein mit der Transkriptions-Aktivierungsdomäne exprimiert wird; die Reihe C mit Zellen bestückt ist, die den pGAD424 Vektor tragen, von dem das Gleitklammergen aus *Archaeoglobus fulgidus* als Fusionsprotein mit der Transkriptions-Aktivierungsdomäne exprimiert wird; die Reihe D nicht mit Zellen bestückt ist und die Reihe E mit Zellen bestückt ist, die den pGAD424 Vektor tragen, von dem das Elongationsproteingen aus *Archaeoglobus fulgidus* als Fusionsprotein mit der Transkriptions-Aktivierungsdomäne exprimiert wird.

- 10 Die Spalte 1 ist mit Zellen bestückt, die den leeren pGBT9 Vektor (Clontech, Palo Alto, USA) tragen; die Spalte 2 ist bestückt mit Zellen, die den pGBT9 Vektor tragen, von dem das *Saccharomyces cerevisiae* Gen UFD3 als Fusionsprotein mit der DNA-Bindedomäne exprimiert wird; die Spalte 3 ist bestückt mit Zellen, die den pGBT9 Vektor tragen, von dem das Gleitklammerprotein aus *Archaeoglobus fulgidus* als Fusionsprotein mit der DNA-Bindedomäne exprimiert wird; die Spalte 4 ist bestückt mit Zellen, die den pGBT9 Vektor tragen, von dem das Kopplungsprotein aus *Archaeoglobus fulgidus* als Fusionsprotein mit der DNA-Bindedomäne exprimiert wird, und
- 20 Spalte 5 ist bestückt mit Zellen, die den pGBT9 Vektor tragen, von dem das Elongationsprotein aus *Archaeoglobus fulgidus* als Fusionsprotein mit der DNA-Bindedomäne exprimiert wird.

### **Beispiel 7: Polymeraseketten-Reaktion von genomischer DNA unter Verwendung eines chimären Elongationsproteins**

Beispiel 7 zeigt den Einfluß des Gleitklammerproteins auf die Effizienz einer PCR – Reaktion an längeren DNA – Fragmenten unter Verwendung des erfindungsgemäßen rekombinanten chimären Proteins. In dieser PCR - Reaktion wurde ein 4954bp großes Fragment ausgehend von humaner genomischer DNA amplifiziert. Als Polymerase wurde das erfindungsgemäße rekombinante chimäre

Protein verwendet. Die Bedingungen entsprechen den Standardbedingungen einer PCR-Reaktion: pH 8,3, 0,4mM jedes dNTP und 20 pmol von jedem Primer in einer Reaktion. Eingesetzt wurde ein erster Primer ((SEQ ID NO.: 15): 5'-AGGAACAACATATGACGCACTCT-3') und ein zweiter Primer ((SEQ ID NO.: 16): (5'-TAGGTGGCCTGCAGTAATGTTAG-3').

Es zeigt sich, daß nur das erfindungsgemäße rekombinante chimäre Protein, in seiner Sequenz dargestellt in Fig. 6A, unter Stimulation durch das Gleitklammerproteins PCNA ein eindeutiges und spezifisches PCR – Produkt generieren ließ, während der gleiche Reaktionsansatz mit nativer Taq DNA-Polymerase kein eindeutiges Produkt ergab. Das erklärt sich durch die Interaktion des Gleitklammerproteins mit dem erfindungsgemäß rekombinanten chimären Protein und einer damit verbundenen geringeren Dissoziation des erfindungsgemäßen rekombinanten chimären Proteins von der Ziel- DNA.

Im folgenden werden die Figuren detaillierter besprochen, auf die in den vorstehenden Beispielen zum Teil bereits Bezug genommen wurde.

Sämtliche Sequenzalignments, wie sie hierin in den folgenden Figuren offenbart werden, wurden mit dem BLAST Algorithmus nach Altschul, S.F., Gish, W. Miller, W., Myers, E.W., and Lipman, D.J., J. Mol. Biol. 215, 403-410 (1990) erstellt.

### Fig. 1:

Fig. 1 zeigt Sequenzalignments von insgesamt vier verschiedenen Elongationsproteindomänen aus verschiedenen Organismen, so für das Elongationsprotein 1 aus Mensch, *Archaeoglobus fulgidus*, *Methanococcus thermoautotrophicus*, PHBT (*Pyrococcus horikoshii*), und *Methanococcus janashii*. Direkt unter der Alignierung steht bei den Figuren häufig eine Konsensussequenz in einer Schreibweise, bei der variierende Positionen durch ein einzelnes Zeichen angegeben werden. Die in eckigen

Klammern angegeben Aminosäuren stellen jene Aminosäuren, ausgedrückt im Einbuchstabencode, dar, die an der speziellen Position stehen können.

Die Aminosäuren werden hierbei gemäß der Standard –IUPAC-Einbuchstaben-Nomenklatur benannt und gemäß dem Prosite Pattern- Beschreibungsstandard aufgeführt. Dabei sind die folgenden Aminosäuregruppen häufig zusammengefaßt:

G,A,V,L,I,M,P,F oder W (Aminosäuren mit nicht polaren Seitenketten), direkt unter der Sequenz oftmals auch als "\$" geschrieben,

S,T,N,Q,Y, oder C (Aminosäure mit ungeladenen polaren Seitenketten),

10 K,R,H,D oder E (Aminosäure mit geladenen und polaren Seitenketten) direkt unter der Sequenz oftmals auch als "&" geschrieben,

außerdem bedeutet X in den Sequenzen bzw. Sequenzprotokollen jede beliebige Aminosäure oder Insertion oder Deletion

Fig. 1 gibt vier verschiedene Konsensussequenzen für verschiedene Bereiche von Elongationsproteinen an, wobei z.B. das Elongationsprotein 3 häufig in Eubakterien gefunden werden kann, das Elongationsprotein 4 auch die Taq-Polymerase umfaßt und häufig in Pol I-Typ-Polymerasen gefunden werden kann. Elongationsprotein 3 zeigt somit ein Alignment einer konservierten Regionen des Elongationsproteins aus  
20 Eubakterien, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: DP3A\_ECOLI: DNA Pol III, alpha subunit, Escherichia coli, BB0579: DNA Pol III, alpha subunit, Borrelia burgdorferi, DP3A\_HELPY: DNA Pol III, alpha subunit, Helicobacter pylori AA50: Aquifex aeolicus, section 50 und DP3A\_SALTY: DNA Pol III, alpha subunit, Salmonella typhimurium).

### Fig. 2:

Fig. 2 zeigt die Skizze einer möglichen Form des erfindungsgemäßen rekombinanten chimären Proteins, wobei die Gleitkammer als die Syntheseaktivität der Polymerase  
30 erhöhender Faktor über die Domäne mit Wechselwirkung-bedingender Sequenz an das die Nukleinsäuresynthese-aktivität tragende Elongationsprotein bindet.

**Fig. 3:**

Fig. 3 zeigt vier Gleitklammerkonsensussequenzen

Folgende Gene sind für Gleitklammer Region 1 gezeigt: Humanes PCNA, sowie Orthologe dazu, aus *Archaeoglobus fulgidus*, aus *Methanococcus janashii*, aus *Pyrococcus horikoshii* und aus *Methanococcus thermoautotrophicus*.

- 10 Region 2 zeigt ein Alignment einer zweiten konservierten Regionen der Gleitklammer aus Eukaryonten und Archae, sowie die daraus hergeleiteten Konsensussequenzen. Das Alignment wurde dabei unter Verwendung einer zweiten Region der oben bei Region 1 bereits verwendeten Gleitklammern erstellt.

- Region 3 zeigt ein Alignment einer konservierten Regionen der Gleitklammer aus Eubakterien, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: AAPOL3B, DP3B\_ECOLI, S.TYPHIM, DP3B\_PROMI, DP3B\_PSEPU und DP3B\_STRCO (AAPOL3B: Aquifex Aeolicus sektion 93: DP3B\_ECOLI : DNA Pol III, beta chain, Escherichia coli, S.TYPHIM: DNA Pol III, beta chain, Salmonella typhimurium, P3B\_PROMI : DNA Pol III, beta chain, Proteus mirabilis DP3B\_PSEPU: DNA Pol III, beta chain, Pseudomonas putida DP3B\_STRCO: DNA Pol III, beta chain, Streptomyces coelicolor).
- 20

Region 4 zeigt ein Alignment einer zweiten konservierten Regionen der Gleitklammer aus Eubakterien (siehe Organismen in Region 3), sowie die daraus hergeleiteten Konsensussequenzen

**Fig. 4:**

- 30 Fig. 4 zeigt die Interaktion von chimären Proteinen im Hefe-Zweihybridsystem. Das in Fig. 4 dargestellte Ergebnis beweist, daß die Eigenschaft eine Wechselwirkung mit



dem Gleitklammerprotein Af0335 zu bedingen, auf ein anders Protein, insbesondere auf eine andere Polymerase, verpflanzt werden kann, um dort eine Wechselwirkung mit dem Gleitklammerprotein hervorzurufen. Es ist gezeigt, daß das GKP sich selbst bindet und das GKP eine Taqfusionsprotein bindet, welches eine Wechselwirkung-bedingende Sequenz aus *Archaeoglobus fulgidus* umfaßt.

**Fig. 5:**

10 Fig. 5 zeigt drei Alignments konservierter Regionen Wechselwirkung-bedingender Sequenz aus verschiedenen Organismen und für verschiedene Gene:

Sequenz 1 (Polymerasegen, Af 0497 Homologe): *Archaeoglobus fulgidus*, *Pyrodictium occultum*, *Aeropyrum pernix*, *Pyrococcus glycovorans*, *Pyrococcus furiosus*, *Thermococcus gorgonarius*, *Pyrococcus abyssi*, *Pyrococcus horikoshii*, *Thermococcus litoralis*, *Thermococcus fumicolans* und *Methanococcus jannaschii*

Sequenz 2 (Polymerasegen, Af 1722 Homologe), *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Pyrococcus furiosus*, *Methanobacterium thermoautotrophicum*, *Pyrococcus horikoshii* und *Pyrococcus abyssi* und

20

Sequenz 3 (Af 1347 Homologe), *Archaeoglobus fulgidus*, *Pyrococcus abyssi*, *Pyrococcus horikoshii*, *Methanobacterium thermoautotrophicum* und *Aeropyrum pernix*.

Die Wechselwirkung-bedingende Sequenzen können in ein erfindungsgemäßes chimaäres Protein eingebracht werden um ein GKP zu binden.

**Fig. 5B:**

30 Fig. 5B zeigt zwei Alignments konservierter Regionen Wechselwirkung-bedingender Sequenz aus verschiedenen Organismen und für verschiedene Gene:

Sequenz 4 (RNase-Gen, Af 0621 Homologe): *Archaeoglobus fulgidus*, *Pyrococcus abyssi*, *Pyrococcus horikoshii* und *Arabidopsis thaliana* und

Sequenz 5 (Polymerase-Gen): *Archaeoglobus fulgidus*, *Pyrococcus abyssi*, *Methanococcus thermoautotrophicus*, *Methanococcus janashii*, *Mus musculus* und *Homo sapiens*.

10 Die Wechselwirkung-bedingende Sequenzen können in ein erfindungsgemäßes chimäres Protein eingebracht werden um ein GKP zu binden.

#### Fig. 5C:

Fig. 5C zeigt eine Liste von Charakteren die verwendet werden um ein oder mehrere verschiedene Aminosäuren zu Kennzeichnen, wobei unter "Klasse" die Aminosäure oder die Eigenschaft der Gruppe steht, unter "key" das verwendete Zeichen und unter "Rest" die Aminosäuren die in der Gruppe sind.

#### Fig. 6A:

20

Die Fig. 6A zeigt die Gesamtsequenz einer bevorzugten Ausführungsform des erfindungsgemäßen rekombinanten chimären Proteins. Hier stammt das Elongationsprotein und damit die Nukleinsäuresyntheseaktivität aus *Thermus aquaticus* und die Wechselwirkung-bedingende Sequenz aus *Archaeoglobus fulgidus*.

#### Fig. 6B:

Die Fig. 6B zeigt eine graphische Übersicht des Aufbaus einer bevorzugten Ausführungsform des erfindungsgemäßen rekombinanten chimären Proteins. Hier ist das

Elongationsprotein aus *Thermus aquaticus*. und die wechselwirkungsbedingende Sequenz aus *Archaeoglobus fulgidus*.

**Fig. 7:**

Fig. 7 zeigt die Verwendung eines erfindungsgemäßen rekombinanten chimären Proteins in der PCR, dessen Sequenz in Fig. 6A abgebildet ist. In allen Spuren dieses 1% Agarose Gels ist jeweils eine Bande für das Amplifikat einer PCR – Reaktion ausgehend von Plasmid-DNA (pCR 2.1 Vektor inklusive eines 463bp Fragments; Invitrogen, 9704 CH Groningen; Netherlands), die unter folgenden Bedingungen (95°C 5' Denaturierung; 30 x {95°C 30'' Denaturierung; 55°C 30'' Hybridisierung; 72°C 40'' Elongation}; 72°C 7 zusätzliche Elongation) durchgeführt wurde, mit einer Größe von 463 bp zu erkennen. Die Reaktionen wurden zu je 20µl eines 50µl Ansatzes auf das Gel aufgetragen. Spur 7 zeigt einen Größenstandard.

In den Spuren 2 (0,3µg PCNA) und 3 (2,4µg PCNA) kann man erkennen, daß sich die Intensität der Bande nach Zugabe steigender Mengen Gleitklammerprotein (PCNA) als stimulierendes Agens bei gleichbleibender  $Mg^{2+}$  - Ionenkonzentration (2,5mM) gegenüber der Spur 1 (PCR-Reaktion ohne PCNA) erhöht. Eine weitere Steigerung der Intensität der Bande und damit verbunden eine erhöhte Ausbeute der PCR-Reaktion kann man bei einer höheren  $Mg^{2+}$  - Ionenkonzentration (3,5mM) erreichen, siehe Spur 4 (ohne PCNA) und Spuren 5 (0,3µg PCNA) und 6 (2,4µg PCNA) bei zunehmender Menge an PCNA.

**Fig. 8A:**

Graphische Darstellung aller Fragmente von Af0497 (siehe oben) , die mit dem Gleitklammerprotein aus *Archaeoglobus fulgidus* (Af0335) interagieren. Mit "B" gekennzeichnete Pfeile repräsentieren Klone, die als Fusion mit der DNA-Bindedomäne

gefunden wurden, Mit "A" gekennzeichnete Pfeile repräsentieren Klone, die als Fusion mit der Aktivierungsdomäne gefunden wurden.

### Fig. 8B:

Fig. 8B zeigt ein Alignment der C-terminalen Sequenzen der folgenden Gene Af 0497-grUE, Af 1195 Rfc, Af 0264 Rad2-Fen1, Af 0621 RNaseH, , *Archaeoglobus fulgidus*.

### 10 Fig. 9:

Fig. 9 zeigt das Ergebnis eines Hefe-Zweihybridtests und belegt eine Bindung zwischen den Hefeproteinen ufd3 und cdc48 (2B; Positivkontrolle), zwischen dem Elongationsprotein und dem Gleitklammerprotein (3E; und in der umgekehrten Orientierung 5C), zwischen dem Gleitklammerprotein und dem Gleitklammerprotein (3C) und zwischen dem Kopplungsprotein und dem Gleitklammerprotein (4C). Die Fusionsproteine werden exprimiert von den Vektoren pGBT9 (Vektor) in Spalte 1, pGBT9::ufd3 (Positivkontrolle) in 2, pGBT9::PCNA in 3, pGBT9::klUE in 4, pGBT9::grUE in 5, pGAD424 (Vektor) in Reihe A, pGAD424::cdc48 in B (Positivkontrolle), pGAD424::PCNA in C, leer in D und pGAD::grUE in E.

Interaktionen von Proteinen aus *Archaeoglobus fulgidus* wurden mit dem Y2H System gezeigt. Die kodierenden Bereiche von Genen aus *Archaeoglobus fulgidus*, deren Genprodukte in vitro verwendeten werden können, wurden mittels PCR amplifiziert, in die Vektoren pGBT9 (vertikale Spalten der Fig. 27) und pGAD424 (horizontale Reihen der Fig. 9) kloniert und durch gap-repair in Hefe PJ69-4a (für pGAD424) und PJ69-4alpha (für pGBT9) als Hybridproteine zur Expression gebracht. Ebenso wurde eine positive Kontrolle mittels PCR amplifiziert, in die Vektoren pGBT9 (siehe auch vertikale Spalten der Fig. 9) und pGAD424 (siehe auch horizontale Reihen der Fig. 9) kloniert und durch gap-repair in den Hefestamm PJ69-4a (für pGAD424) und

PJ69-4alpha (für pGBT9) als Hybridproteine zur Expression gebracht. Diploide Zellen, die beide Vektoren enthalten, wurden durch Paarung entsprechend dem gezeigten Raster in Fig. 9 erzeugt. Die Expression von drei unabhängigen Reportern (*HIS3*, *ADE2* und *MEL1*) wurde gemessen. Gezeigt ist in Fig 9 das Wachstum auf Medium ohne Histidin und Adenin. In diesem Experiment wachsen jene Zellen, die beide Vektoren tragen und zudem die Expressionsprodukte dieser beiden Vektoren einander binden. Was dadurch bedingt wird, daß in Folge der Bindung der Expressionsprodukte Transkription initiiert wird, die dazu führt, daß die Histidin und Adenin Auxotrophie aufgehoben wird.

10

Alle hier positiven Klone waren auch positiv bezüglich der Expression des *MEL1* Gens. Die Handhabung des two-hybrid systems erfolgte gemäß der Anleitung der Firma Clontech (yeast protocols handbook, PT3024-1).

### Fig. 10

In allen Spuren dieses 1% Agarose Gels mit Ausnahme von Spur 5, die einen Größenstandard zeigt, ist das Amplifikat einer PCR – Reaktion ausgehend von humaner genomischer DNA mit einer Größe von 4954 Basenpaaren (Roche Diagnostics, D – 68305 Mannheim), die unter den folgenden Bedingungen: 95°C 5' Denaturierung; 35 x {95°C 30'' Denaturierung; 47°C 30'' Hybridisierung; 72°C 12' Elongation}; 72°C 10 zusätzliche Elongation, durchgeführt wurde, zu je 20µl eines 50µl Ansatzes auf das Gel aufgetragen worden. In Spur 3 (2,4µg PCNA) kann man erkennen, daß man eine Bande nach Zugabe steigender Mengen des Gleitklammerproteins (PCNA aus *Archaeoglobus fulgidus*) zum erfindungsgemäßen rekombinanten chimären Proteinkomplexe als stimulierendes Agens bei gleichbleibender  $Mg^{2+}$  - Ionenkonzentration (6mM) gegenüber der Spur 1 (PCR-Reaktion ohne PCNA) und Spur 2 (0,3µg PCNA) erhält, während sich hingegen bei gleicher  $Mg^{2+}$  - Ionenkonzentration mit Taq DNA-Polymerase (Spur 4) kein eindeutiges Produkt in der PCR - Reaktion generieren ließ.

30



Die in der vorliegenden Beschreibung sowie in den Ansprüchen offenbarten Merkmale der Erfindung können sowohl einzeln als auch in beliebiger Kombination für die Verwirklichung der Erfindung in ihren verschiedenen Ausführungsformen wesentlich sein.

Ansprüche

48

## 1. Rekombinantes Chimäres Protein umfassend

- a) eine erste Domäne mit Nucleinsäuresyntheseaktivität, und
- b) eine Wechselwirkung vermittelnde Sequenz,

dadurch gekennzeichnet, daß durch die Wechselwirkung vermittelnde Sequenz ein Komplex aus Nukleinsäuresyntheseaktivität und Gleitklammerprotein gebildet wird, wobei der Komplex verschieden ist von dem Komplex, den die Nukleinsäuresyntheseaktivität und/oder das Gleitklammerprotein mit ihren natürlichen Wechselwirkungspartner (n) ausbildet.

## 2. Chimäres Protein umfassend

einen Nukleinsäuresyntheseaktivität aufweisenden Anteil, wobei der Anteil aus einem einen Nukleinsäuresyntheseaktivität aufweisenden Anteil und einen Wechselwirkung vermittelnden Anteil umfassenden Basisprotein stammt, und

mindestens einen eine Wechselwirkung vermittelnden Anteil,

wobei der Wechselwirkung vermittelnde Anteil verschieden ist von dem Wechselwirkung vermittelnden Anteil des Basisproteins.

## 3. Chimäres Protein nach Anspruch 2, dadurch gekennzeichnet, daß das Protein den Wechselwirkung vermittelnden Anteil des Basisproteins umfaßt.

#### 4. Chimäres Protein umfassend

einen Nukleinsäuresyntheseaktivität aufweisenden Anteil, wobei der Anteil aus einem Basisprotein stammt, das einen Nukleinsäuresyntheseaktivität aufweisenden Anteil aber keinen Wechselwirkung vermittelnden Anteil umfaßt, und

einen Wechselwirkung vermittelnden Anteil.

5. Chimäres Protein nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß der Wechselwirkung vermittelnde Anteil eine Bindung zwischen der Nukleinsäuresyntheseaktivität und einem die Syntheseleistung der Nukleinsäuresyntheseaktivität beeinflussenden Faktor vermittelt.

6. Chimäres Protein nach Anspruch 5, dadurch gekennzeichnet, daß der Faktor ein Gleitklammer-protein ist.

7. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß

die Nucleinsäuresyntheseaktivität eine Konsensuspeptidsequenzen umfaßt, wobei die Sequenz ausgewählt ist aus der Gruppe, die die Sequenzen SEQ ID NO.: 1, 2 und 3 umfaßt.

8. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, daß

das Gleitklammerprotein eine Konsensuspeptidsequenz umfaßt, wobei die Sequenz ausgewählt ist aus der Gruppe, die die Sequenzen SEQ ID NO.: 4, 5, 6 und 7 umfaßt .

9. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß

die Wechselwirkung vermittelnde Sequenz eine Konsensuspeptidsequenz umfaßt, welche ausgewählt ist aus der Gruppe, die die Sequenzen SEQ ID NO.: 8, 9, 10 11 und 12 umfaßt.

10. Rekombinantes chimäres Protein nach Anspruch 9, dadurch gekennzeichnet, daß

die Wechselwirkung vermittelnde Sequenz eine Konsensuspeptidsequenz gemäß SEQ ID NO.: 8 umfaßt.

11. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß

die Wechselwirkung vermittelnde Sequenz am C-terminalen Ende der die Nucleinsäuresyntheseaktivität tragenden Sequenz liegt.

12. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 11, dadurch gekennzeichnet, daß

zwischen der Wechselwirkung vermittelnden Sequenz und der die Nukleinsäureaktivität tragenden Sequenz ein Linker angeordnet ist.

13. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 12, dadurch gekennzeichnet, daß

das rekombinante chimäre Protein thermostabil ist.

14. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 13, dadurch gekennzeichnet, daß

das Protein eine DNA Polymerase-Aktivität aufweist.

15. Rekombinantes chimäres Protein nach Anspruch 14, dadurch gekennzeichnet, daß

das Protein eine 3'-5' Exonukleaseaktivität aufweist.

16. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 15, dadurch gekennzeichnet, daß

das Protein eine RNA Polymerase-Aktivität aufweist.

17. Rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 14, dadurch gekennzeichnet, daß

das Protein eine reverse Transkriptase-Aktivität aufweist.

18. Rekombinantes chimäres Protein gemäß Anspruch 14, dadurch gekennzeichnet, daß

sich die Einbaurate von dNTPs und ddNTPs durch die Nukleinsäuresyntheseaktivität um einen Faktor von weniger als 5 unterscheidet.



19. Komplex umfassend

- a) ein rekombinantes chimäres Protein gemäß einem der Ansprüche 1 bis 18  
und
- b) ein Gleitklammerprotein.

20. Komplex gemäß Anspruch 19,

wobei das Gleitklammerprotein eine Konsensuspeptidsequenz umfaßt, welche ausgewählt ist aus der Gruppe, die die Sequenzen gemäß SEQ ID NO.: 4, 5, 6 und 7 umfaßt.

21. Komplex nach Anspruch 19 oder 20, weiter umfassend eine Nukleinsäure.

22. Nukleinsäure kodierend für ein rekombinantes chimäres Protein gemäß einem der Ansprüche 1 bis 18.

23. Vektor umfassend eine Nukleinsäure gemäß Anspruch 22.

24. Vektor gemäß Anspruch 23, wobei dieser ein Expressionsvektor ist.

25. Zelle umfassend einen Vektor gemäß Anspruch 23 oder 24.

26. Verwendung eines rekombinanten chimären Proteins nach einem der Ansprüche 1 bis 18 zur Elongation von Nukleinsäuren.

27. Verwendung eines rekombinanten chimären Proteins gemäß einem der Ansprüche 1 bis 18 zur Amplifikation von Nukleinsäuren.
28. Verwendung eines rekombinanten chimären Proteins gemäß einem der Ansprüche 1 bis 18 zur reversen Transkription von RNA in DNA.
29. Verwendung eines rekombinanten chimären Proteins gemäß einem der Ansprüche 1 bis 18 zur Sequenzierung von DNA.
30. Reagenzien-Kit zur Elongation und/oder Amplifikation und/oder reversen Transkription und/oder Sequenzierung und/oder Markierung von Nukleinsäuren, umfassend in einem oder mehreren getrennten Behältern
- a) ein rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 18 und/oder
  - b) einen Komplex nach einem der Ansprüche 19 bis 21 und bevorzugterweise
  - c) gegebenenfalls zumindest einen Primer, Puffer, Nukleotide, Kofaktoren und/oder Pyrophosphatase.
31. Kit nach Anspruch 30, dadurch gekennzeichnet,
- daß er zur Amplifikation von Nukleinsäuren neben den Substanzen a) und/oder b), Deoxynukleotide oder/und Derivate davon umfaßt.
32. Kit nach Anspruch 30 oder 31, dadurch gekennzeichnet,
- daß er eine DNA Polymerase mit 3'-5' Exonukleaseaktivität umfaßt.

33. Kit nach Anspruch 30, dadurch gekennzeichnet,

daß er zur Reversen Transkription Substanzen nach a) und/oder b) enthält, welche eine Reverse Transkriptase-Aktivität aufweisen, und bevorzugterweise Deoxynukleotide und/oder Derivate davon.

34. Kit nach Anspruch 30, dadurch gekennzeichnet,

daß er zur Sequenzierung neben Deoxynukleotiden oder Ribonukleotiden oder/und deren Derivaten, Dideoxynukleotide oder/und deren Derivate enthält.

35. Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, wobei die zu elongierende Nukleinsäure oder zumindest ein Strang davon mit mindestens einem Primer unter Hybridisierungsbedingungen versehen wird, wobei der Primer hinreichend komplementär zu einem Teil der oder einem flankierenden Bereich der zu elongierenden Nukleinsäure ist und durch eine Polymerase in Anwesenheit von Nukleotiden eine Primer-Elongation erfolgt, dadurch gekennzeichnet,

daß als Polymerase ein rekombinantes chimäres Protein gemäß einem der Ansprüche 1 bis 18 verwendet wird und

daß bevorzugterweise ein Gleitklammerprotein in der Reaktion vorliegt.

36. Verfahren zur Amplifikation einer Nukleinsäure, wobei die zu amplifizierende Nukleinsäure mit mindestens zwei Primern unter Hybridisierungsbedingungen versetzt wird, wobei ein jeder der beiden Primer jeweils komplementär zu einem Teil der oder einem flankierenden Bereich der zu amplifizierenden Nukleinsäure ist und durch eine Polymerase in Anwesenheit von Nukleotiden eine Primer-Elongation erfolgt, dadurch gekennzeichnet,

daß als Polymerase ein chimäres Protein nach einem der Ansprüche 1 bis 18 verwendet wird und

daß bevorzugterweise ein Gleitklammerprotein der Reaktion zugesetzt wird.

37. Verfahren nach Anspruch 36, dadurch gekennzeichnet,

daß man eine Polymerase-Ketten-Reaktion durchführt.

38. Verfahren nach Anspruch 37, dadurch gekennzeichnet,

daß der Reaktionsansatz zwei DNA Polymerasen umfaßt, von denen mindestens eine eine 3'-5' Exonukleaseaktivität aufweist, wobei die 3'-5'-Exonukleaseaktivität entweder durch das rekombinante chimäre Protein oder durch eine weitere Polymerase dem Reaktionsansatz zugesetzt wird.

39. Verfahren nach Anspruch 38, dadurch gekennzeichnet,

daß zwei rekombinante chimäre Proteine nach einem der Ansprüche 1 bis 18 in der Reaktion vorliegen, wobei eines ein solches nach Anspruch 14 und das andere ein solches nach Anspruch 15 ist.

40. Verfahren nach Anspruch 35, dadurch gekennzeichnet,

daß man zur Sequenzierung von Nukleinsäuren ausgehend von einem Primer, der zu einem der zu sequenzierenden Nukleinsäure benachbarten Bereich komplementär ist, eine Template-abhängige Elongation bzw. Reverse-Transkription unter Verwendung von Deoxynukleotiden und Dideoxynukleotiden oder deren jeweiligen Derivaten gemäß der Methode von Sanger durchführt.

41. Verfahren nach Anspruch 35, dadurch gekennzeichnet,

daß man bei der Elongation der Nukleinsäuren Markierungen einfügt.

42. Verfahren nach Anspruch 41, dadurch gekennzeichnet, daß ein Agens verwendet wird, das ausgewählt ist aus der Gruppe, die markierte Primer, markierte Deoxynukleotide und Derivate davon, markierte Dideoxynukleotide und Derivate davon und markierte Ribonukleotide und Derivate davon umfaßt.

43. Verfahren zur Markierung von Nukleinsäuren durch Erzeugung einzelner Brüche in Phosphodiesterbindungen der Nukleinsäurekette und Ersatz eines Nukleotids an den Bruchstellen durch ein markiertes Nukleotid mit Hilfe einer Polymerase, dadurch gekennzeichnet,

daß als Polymerase ein rekombinantes chimäres Protein nach einem der Ansprüche 1 bis 18 eingesetzt wird.

44. Antikörper gerichtet gegen ein chimäres Protein nach einem der Ansprüche 1 bis 18.



Fig. 1

1/13

## Elongationprotein 1

DPOD\_HUMAN : D V I T G Y N I Q N F D L P Y L I S R A  
 DPOL\_ARCFU : D I I V G Y N Q D A F D W P Y L R K R A  
 MTH1208 : D I L V G Y N S D N F D F P Y I T R R A  
 PHBT047 : D V I I T Y N G D N F D F P Y L L K R A  
 MJ0885 : D V I Y T Y N G D N F D F P Y L K A R A  
 Konsensus : D \$ \$ X X Y N X X X F D X P Y \$ X X R A  
 SEQ ID NO1 : D-[GAVLIMPFW]-[GAVLIMPFW]-X-X-Y-N-X-X-X-F-D-X-P-Y-[GAVLIMPFW]-X-X-R-A

## Elongationsprotein 2

AF1722 : A V R T A V A I M T E G V V A A P I E G I A R V R I  
 MJ1630 : A V R T A L A V L T E G I V A A P L E G I A D V K I  
 PfuORF3 : A V R T A L A I L T E G I V S A P L E G I A D V K I  
 MTH1536 : A L R T A L A I L T E G V V A A P L E G I A R V R I  
 PHBN021 : A V R T A L A I L T E G V V S A P I E G I A S V K I  
 Konsensus: A \$ R T A \$ A \$ \$ T E G \$ V X A P \$ E G I A X V & I  
 SEQ ID NO.2: A-[GAVLIMPFW]-R-T-A-[GAVLIMPFW]-A-[GAVLIMPFW]-[GAVLIMPFW]-T-E-G-  
 [GAVLIMPFW]-V-X-A-P-[GAVLIMPFW]-E-G-I-A-X-V-[KRHDE]-I

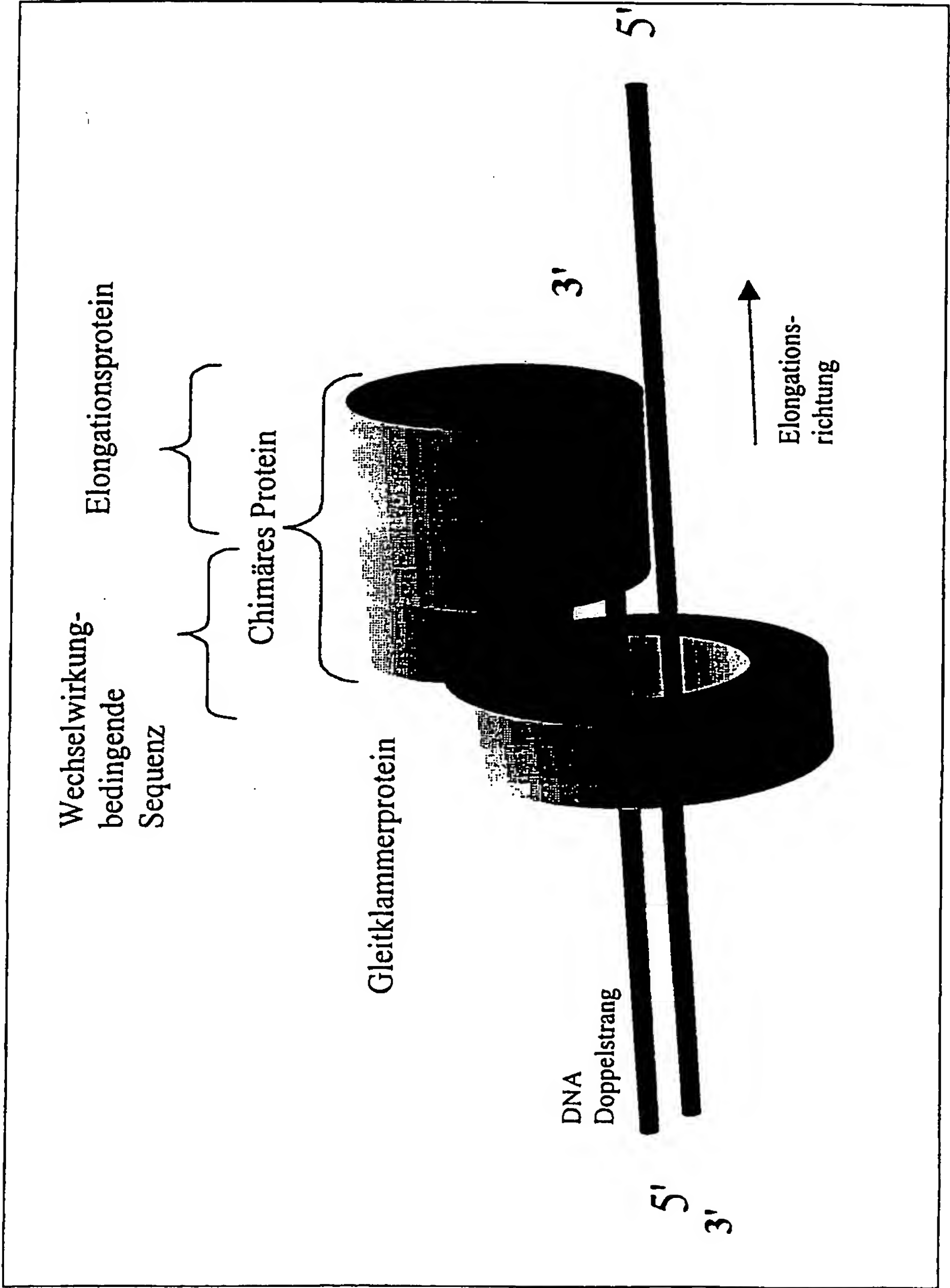
## Elongationsprotein 3 (Eubakterien):

DP3A\_ECOLI : VPVGPGRGSGAGSLVAYALKITDLDPLEFDLLFERFLNPERVSMMPD  
 DP3A\_SALTY : VPVGPGRGSGAGSLVAYALKITDLDPLEFDLLFERFLNPERVSMMPD  
 BB0579 : IPVGAGRGSGAGSIVAYALRITDIDPLKYNLLFERFLNPERISMMPD  
 DP3A\_HELPY : IPVGPGRGSAAGSLVAFALKITDIDPLKYDLLFERFLNPERISMMPD  
 AA50 : IPVGPGRGSAAGSLVAYAIGITDVDPIKHGFLFERFLNPERVSMMPD  
 Konsensus : \$PVG\$GRGSX\$G\$SVAXA\$XITD\$DP\$XXX\$LFERFLNPER\$SMMPD  
 Or written as: [GAVLIMPFW]-P-V-G-[GAVLIMPFW]-G-R-G-S-X-[GAVLIMPFW]-G-S-[GAVLIMPFW]-V-  
 A-X-A-[GAVLIMPFW]-X-I-T-D-[GAVLIMPFW]-D-P-[GAVLIMPFW]-X-X-X-[GAVLIMPFW]-L-F-E-R-  
 F-L-N-P-E-R-[GAVLIMPFW]-S-M-P-D

## Elongationsprotein 4 (Inklusive Taq):

Pol 1: M Q V H D E L  
 T5 : M L V H D S V  
 T7 : A W V H D E I  
 Spn : L Q V H D E I  
 TAQ : L Q V H D E L  
 Spo2 : M H V H D E A  
 Konsensus/SEQ ID NO. 3: [MALM][QLVH]V H D [SE][ALVI]

Fig. 2:



3/13

Fig 3

## Gleitklammer Region 1:

PCNA\_HUMAN : M D S S H V S L V Q L T L R S E G F D T Y R C D  
 PCNA\_METJA : M D P S H V A L V S L E I P R L A F E E Y E A D  
 MTH1312 : L D R S H I T Y V H L E L K A E L F D E Y V C D  
 PHLA008 : M D P S R V V L I D L N L P S S I F S K Y E V D  
 AF0335 : V D P A N V A M V I V D I P K D S F E V Y N I D

Konsensus : \$ D X X X \$ X X \$ X \$ X \$ X X X X F X X Y X X D

SEQ ID NO. 4.: [GAVLIMPFW]-D-X-X-X-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-[GAVLIMPFW]-X-[GAVLIMPFW]-X-X-X-X-F-X-X-Y-X-X-D

## Gleitklammer Region 2:

PCNA\_HUMAN : L K Y Y L A P K I E  
 PCNA\_METJA : L T F L L A P R I E  
 MTH1312 : L S F L L A P R I E  
 PHLA008 : L T F L L A P R V E  
 AF0335 : V E Y I L A P R I E

Konsensus : \$ X X X L A P & \$ E

SEQ ID NO. 5 [GAVLIMPFW]-X(3)-L-A-P-[KRHDE]-[GAVLIMPFW]-E

## Gleitklammer (Eubakterien) Region 3:

AAPOL3B : K L V I T G G K S T Y K L P T A P A E D F P  
 DP3B\_ECOLI : R M L V R S G R S R F S L S T L P A A D F P  
 S.TYPHIM. : R M L V R S G R S R F S L S T L P A A D F P  
 DP3B\_PROMI : R L L V R S G R S R F S L S T L P A S D F P  
 DP3B\_PSEPU : K L L V K A G R S R F T L S T L P A N D F P  
 DP3B\_STRCO : R A T V V C G S S R F T L H T L P V E E Y P

Konsensus : & X X \$ X X G X S X X X L X T \$ P \$ X & X P

SEQ ID NO 6. : [KRHDE]-X-X-[GAVLIMPFW]-X-X-G-X-S-X-X-X-L-X-T-[GAVLIMPFW]-P-[GAVLIMPFW]-X-[KRHDE]-X-P

## Gleitklammer (Eubakterien) Region 4:

AAPOL3B : I I M P M R V  
 DP3B\_ECOLI : V V M P M R L  
 S.TYPHIM. : V V M P M R L  
 DP3B\_PROMI : V V M P M R L  
 DP3B\_PSEPU : V V M P M R L  
 DP3B\_STRCO : L I M P V R L

Konsensus : \$ \$ M P \$ R \$

SEQ ID NO. 7: [GAVLIMPFW]-[GAVLIMPFW]-M-P-[GAVLIMPFW]-R-[GAVLIMPFW]

Fig. 4

	PGBDU	PGBDU, PCNA	PGBDU, Isu	PGBDUTAQ ID ct	PGBDU, ssu
P GAD	-	-	-	-	-
P GAD, PCNA	-	+	+	+	-

+ steht für positive Interaktion

- steht für keine Interaktion

Fig. 5:

5/13

Konsensus Wechselwirkung-bedingende Sequenz 1 (Af0497 grosse Ue):

1.	swiss O29753 DPOL_ARCFU	ERFGYTEASLKGSS..	<u>QMSLDSFFS</u>
2.	trembl D38574 PODNAPB_1	SYFGVTEKQLKAAATV	QRS L F D F F A
3.	swiss O93746 DPO2_AERPE	QYFGVTEKRLKGGGR.	QST L L D F M R
4.	tremblnew AJ004834 PSPAJ4834_1	EAFGYRKEDLRWQKT	KQTGLTAWLN
5.	swiss P80061 DPOL_PYRFU	EGFGYRKEDLR Y Q K T R	QVGLTSWLN
6.	swiss P56689 DPOL_THEGO	RAFGYRKEDLR Y Q K T R	QVGLGAWLK
7.	swiss P77916 DPOL_PYRAB	RAFGYRKEDLR Y Q K T K	QVGLGAWLK
8.	swiss O59610 DPOL_PYRHO	KAFGYRKEDLRWQKT	KQVGLGAWIK
9.	swiss P30317 DPOL_THELI	EAFGYRKEDLR Y Q S S K	QTGLDAWLK
10.	swiss P74918 DPOL_THEFM	KAFGYKKEDLR Y Q K T R	QVGLGAWLK
11.	swiss Q58295 DPOL_METJA	EAVGVSKNELKKEGA.	QLTLDKFFK
12.	KONSENSUS 1	p h h g h p c t p L + h t t X X <u>Q X s L X t a h t</u>	
13.	KONSENSUS 2	SEQ ID NO. 8	<u>Q X [STG] L X [SDAK] [FW] [FML] [SARNK]</u>
14.	Taqfusion	SEQ ID NO. 9	<u>Q M S L D S F F S</u>

- 1 = Archaeoglobus fulgidus
- 2 = Pyrodictium occultum
- 3 = Aeropyrum pernix
- 4 = Pyrococcus glycovorans
- 5 = Pyrococcus furiosus
- 6 = Thermococcus gorgonarius
- 7 = Pyrococcus abyssi
- 8 = Pyrococcus horikoshii
- 9 = Thermococcus litoralis
- 10 = Thermococcus fumicolans
- 11 = Methanococcus jannaschii

Konsensus Wechselwirkung-bedingende Sequenz 2 (Af1722 -DP2)

1.	trembl AE000984 AE000984_2	QVSISDF
2.	trembl U67603 MJU67603_3	QVKLSDF
3.	trembl D84670 D84670_3	QLTLIVD
4.	trembl AE000913 AE000913_9	QSSLDVF
5.	tremblnew AP000001 AP000001_12	QLTLLVN
6.	trembl AJ248283 CNXPAX01_120	QLTLIVN

KONSENSUS/SEQ ID NO. 10: = Q[LSV] p 1 [LIDS] [VD] [FND]

- 1 = Archaeoglobus fulgidus
- 2 = Methanococcus jannaschii
- 3 = Pyrococcus furiosus
- 4 = Methanobacterium thermoautotrophicum
- 5 = Pyrococcus horikoshii
- 6 = Pyrococcus abyssi

Konsensus für Wechselwirkung-bedingende Sequenz 3 (Af1347):

1.	trembl AE001011 AE001011_16	WRALVS...R	QTS LASFFGF ..
2.	trembl AJ248284 CNXPAX02_129	WLIRQL...N	QITIMSFLGV ..
3.	tremblnew AP000007 AP000007_17	WILKQL...N	QRSITSFLGV RS
4.	trembl AE000993 AE000993_12	TLVKRG...R	QRTLWEFM... ..
5.	trembl AE000851 AE000851_5	NLLKELEVSR	QTTLDSFFR. ..



6/13

Fig. 5 (Fortsetzung)

## KONSENSUS

XhhhXXXXXp Q.ol.pahXX XX

- 1 = Archaeoglobus fulgidus
- 2 = Pyrococcus abyssi
- 3 = Pyrococcus horikoshii
- 4 = Methanobacterium thermoautotrophicum
- 5 = Aeropyrum pernix

**7/13**

1. swissnew 029634 RNH2_ARCFU	SNLR.....QKTLDDF
2. trembl AJ248284 CNSPAX02_220	KKIEEKFKRA QLTLDNF
3. swissnew 059351 RNH2_PYRHO	RKIEERFRKN QLTLDKF
4. tremblnew AC018722 AC018722_3	THLKGEVEVA <u>WLVFFCF</u>

1 = *Archaeoglobus fulgidus*  
2 = *Pyrococcus abyssi*  
3 = *Pyrococcus horikoshii*  
5 = *Arabidopsis thaliana*

```

O29753|DPOL_ARCFU      : VAYDYNFDPYFYV
P80061|DPOL_PYRFU      : IEHDRTERPYIYA
O27276|DPOL_METTH      : IAHDRSFRPYIYA
Q58295|DPOL_METJA      : ILKDRDFKPYFYV
P52431|DPOD_MOUSE      : CCHIQQFAPYFYT
M81735|HSDNAPOLD_1     : CCHIHGFAFYFYT

```

SEQ ID NO. 12: : [GAVLIMPFW]-[GAVLIMPFW]-X-[GAVLIMPFW]-X-X-X-X-X-X-X-X-  
X-X-X-X-[GAVLIMPFW]-X-X-X-X-X-F-X-P-Y-F-Y-X

1 = Archaeoglobus fulgidus  
2 = Pyrococcus abyssi  
3 = Methanococcus thermoautotrophicus  
5 = Methanococcus janashii  
6.= Mus musculus  
7.= Homo sapiens

Fig 5C:

8/13

Klasse	Key	Rest
A	A	A
C	C	C
D	D	D
E	E	E
F	F	F
G	G	G
H	H	H
I	I	I
K	K	K
L	L	L
M	M	M
N	N	N
P	P	P
Q	Q	Q
R	R	R
S	S	S
T	T	T
V	V	V
W	W	W
Y	Y	Y
Alkohol	o	S, T
Aliphatisch	l	I, L, V
Jedes	X	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y (auch Insertion oder Deletion)
Aromatisch	a	F, H, W, Y
Geladen	c	D, E, H, K, R
Hydrophob	h	A, C, F, G, H, I, K, L, M, R, T, V, W, Y
Negative	-	D, E
Polar	p	C, D, E, H, K, N, Q, R, S, T
Positiv	+	H, K, R
Klein	s	A, C, D, G, N, P, S, T, V
Winzig	u	A, G, S
Turnlike	t	A, C, D, E, G, H, K, N, Q, R, S, T

9/13

Fig.6 A: SEQ ID NO. 17 (Elongationsprotein & Linker & Wechselwirkung-bedingende Sequenz)

Elongationsprotein:

```
MRGMLPLFEP KGRVLLVDGH HLAYRTFHAL KGLTTSRGEP
VQAVYGFAKS LLKALKEDGD AVIVVFDAKA PSFRHEAYGG
YKAGRAPTPE DFPRQLALIK ELVDLLGLAR LEVPGYEADD
VLASLAKKAE KEGYEVRIIT ADKDLYQLLS DRIHVLHPEG
YLITPAWLWE KYGLRPDQWA DYRALTGDES DNLPGVKGIG
EKTARKLLEE WGSLEALLKN LDRLKPAIRE KILAHMDDLK
LSWDLAKVRT DLPLEVDFAK RREPDRERLR AFLERLEFGS
LLHEFGLLS PKALEEAPWP PPEGAFVGFV LSRKEPMWAD
LLALAAARGG RVHRAPEPYK ALRDLKEARG LLAKDLSVLA
LREGLGLPPG DDPMLLAYLL DPSNTTPEGV ARRYGGEWTE
EAGERAALE RLFANLWGRL EGEERLLWLY REVERPLSAV
LAHMEATGVR LDVAYLRALS LEVAEEIARL EAEVFRLAGH
PFNLNSRDQL ERVLFDELGL PAIGKTEKTG KRSTSAAVLE
ALREAHPIVE KILQYRELTK LKSTYIDPLP DLIHPRTGRL
HTRFNQTATA TGRLLSSDPN LQNI PVRTPL GQIRRAFIA
EEGWLLVALD YSQIELRVLA HLSGDENLIR VFQEGRDIHT
ETASWMFGVP REAVDPLMRR AAKTINFGVL YGMSAHRLSQ
ELAIPYEEAQ AFIERYFQSF PKVRAWIEKT LEEGRRRGYV
ETLFGRRRYV PDLEARVKSV REAAERMAFN MPVQGTAADL
MKLAMVKLFP RLEEMGARM LQVHDELVE APKERAEAVA
RLAKEVMGV YPLAVPLEVE VGIGEDWLSA KE
```

Linker:

PPRPRPKSGI EIKKLDKDY IDNQIIPSVL RILERFGYTE ASLKGSS

Wechselwirkung-bedingende Sequenz:

QMSLDSFFS

Fig. 6 B:

minosäuresequenzübergänge in der Taqfusionschimäre:

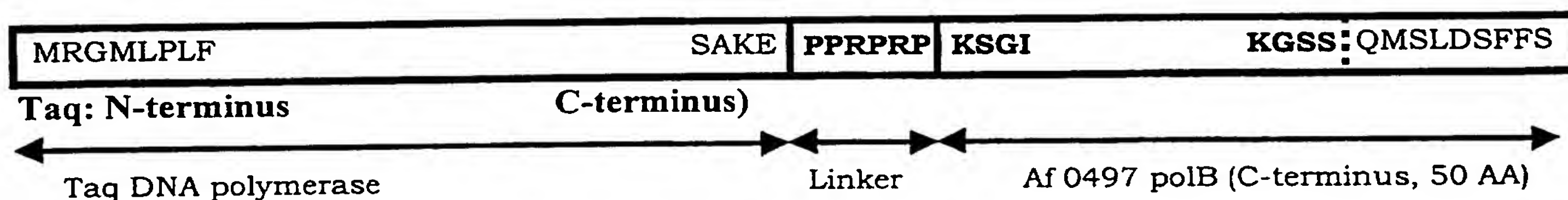
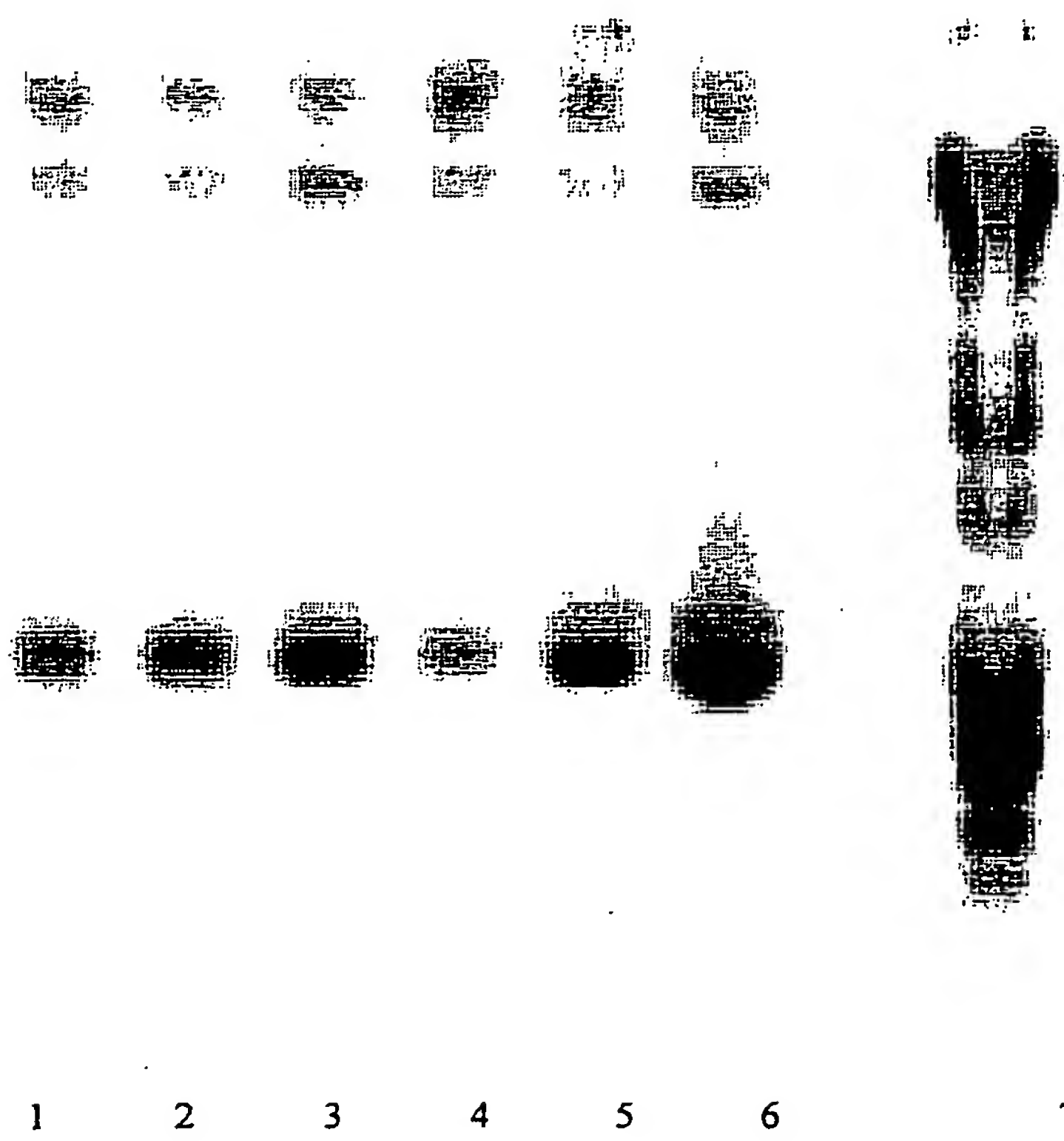


Fig. 7: PCR mit Plasmid DNA (Taqfusionsprotein mit PCNA stimuliert)



Ansatz:

	Konzentrationen	1	2	3	4	5	6
Plasmid DNA	0,5µg/µl	0,7µl	0,7µl	0,7µl	0,7µl	0,7µl	0,7µl
dNTP	25mM each dNTP	0,8µl	0,8µl	0,8µl	0,8µl	0,8µl	0,8µl
Primer fw	20 pmol/µl	1µl	1µl	1µl	1µl	1µl	1µl
Primer rev	20pmol/µl	1µl	1µl	1µl	1µl	1µl	1µl
Puffer	100mM Tris-HCl, 500mM KCl; pH 8,3(20°C)	5µl	5µl	5µl	5µl	5µl	5µl
MgCl <sub>2</sub>	25mM	5µl	5µl	5µl	7µl	7µl	7µl
Storage buffer	50%Glycerol; 100mM Tris-HCl pH 7,9, 100mM KCl; 10mM EDTA	8µl	7µl	0µl	8µl	7µl	0µl
Af PCNA	0,3µg/µl	0µl	1µl	8µl	0µl	1µl	8µl
Taqfusionsprotein		3µl	3µl	3µl	3µl	3µl	3µl
Wasser		25,5µl	25,5µl	25,5µl	23,5µl	23,5µl	23,5µl



Fig. 8 A:

11/13

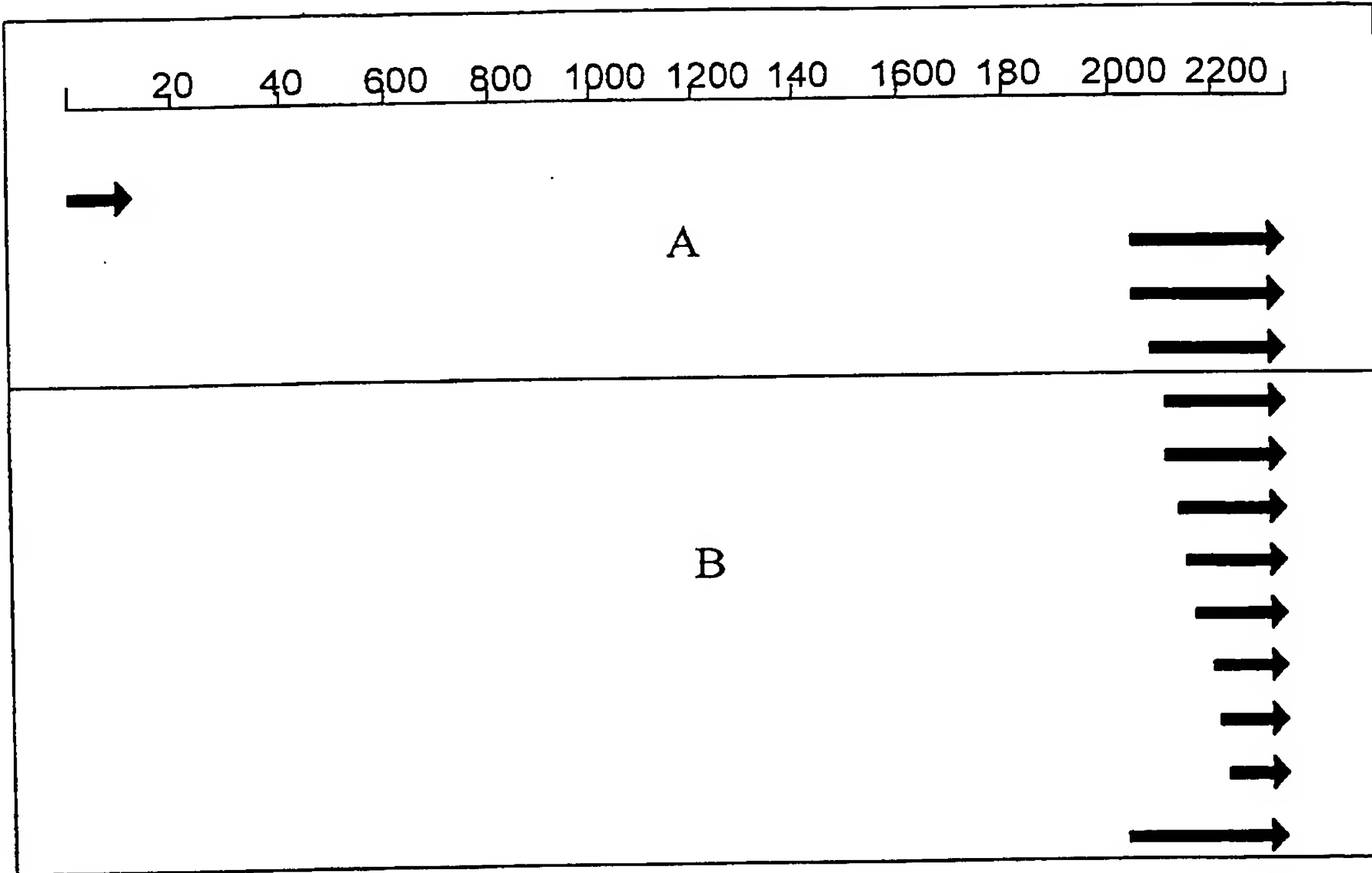


Fig. 8 B:

Label	10	20	30	40	50
AF0497-GrDE	KSGIEIKKLD	KDY-YID-NQ	IIPSVLRILE	RF-GYTEASL	KGSSQMSLDS
AF1195-RFC2	TFLESFVKAE	KEE-KEESVE	ZVAEEKPEEE	R--EEPRAK	KAGKNLTIDS
AF0264-Rad2-Pan1	DYRIEFREPD	FEK-AIEFLC	EEHDFSREVR	EK-ALEKLKA	LKSTQATLER
AF0621-RNaseH	KFGDFGSGYA	SDPRTREVLK	EWIASGRIPS	CV-RMRWKTV	SNLRQKTLD
AF1347-Rad1	--EEGVRKAL	KS--KPEVFE	SFDDALTAAS	RRVENKEWRA	LVSROQTSLAS
AF1552-Rad1	KDTVTLFSLV	KSWDVPTLVD	YYVSVLHLAF	RK-KVEIRQE	EFYGDVEIQK
AF1790-Rad1	GFYRGVFMVN	SST-WQAQTE	FQKKVNLNPM	PG-NVAVYRP	--GGEVIRLR
Label	54				
AF0497-GrDE	FFS-				
AF1195-RFC2	FFS-				
AF0264-Rad2-Pan1	WF--				
AF0621-RNaseH	F---				
AF1347-Rad1	FFGF				
AF1552-Rad1	F---				
AF1790-Rad1	FYGE				

Fig. 9:

12/13

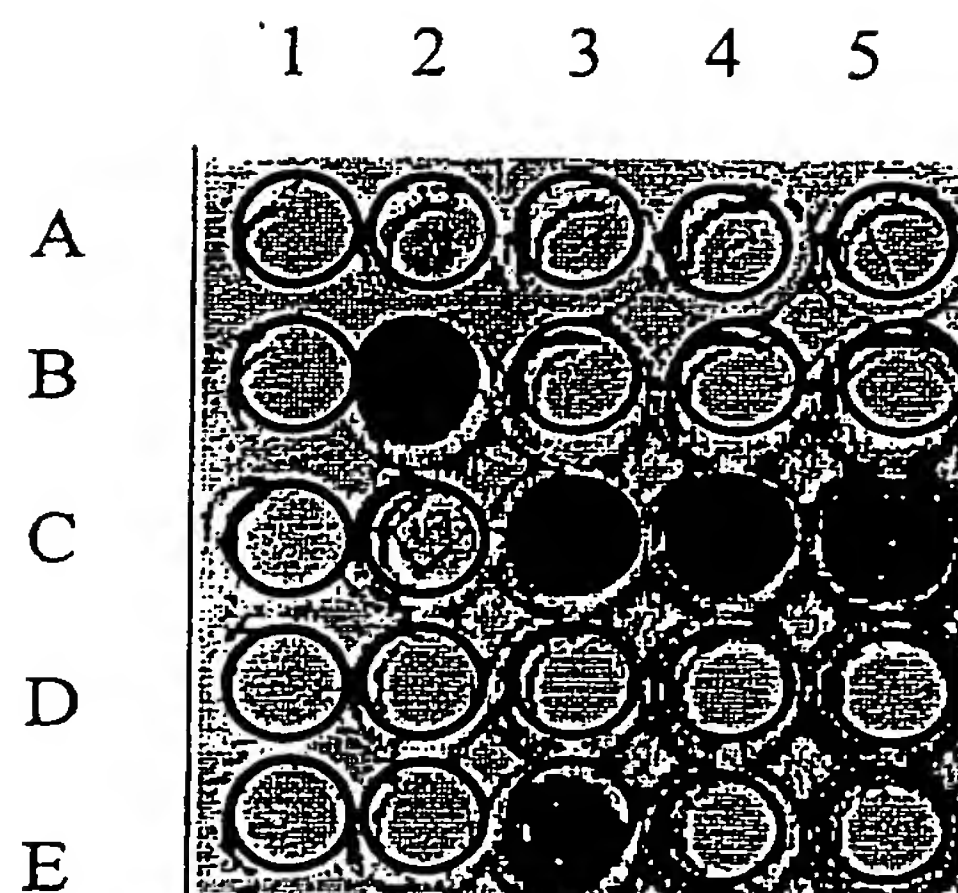
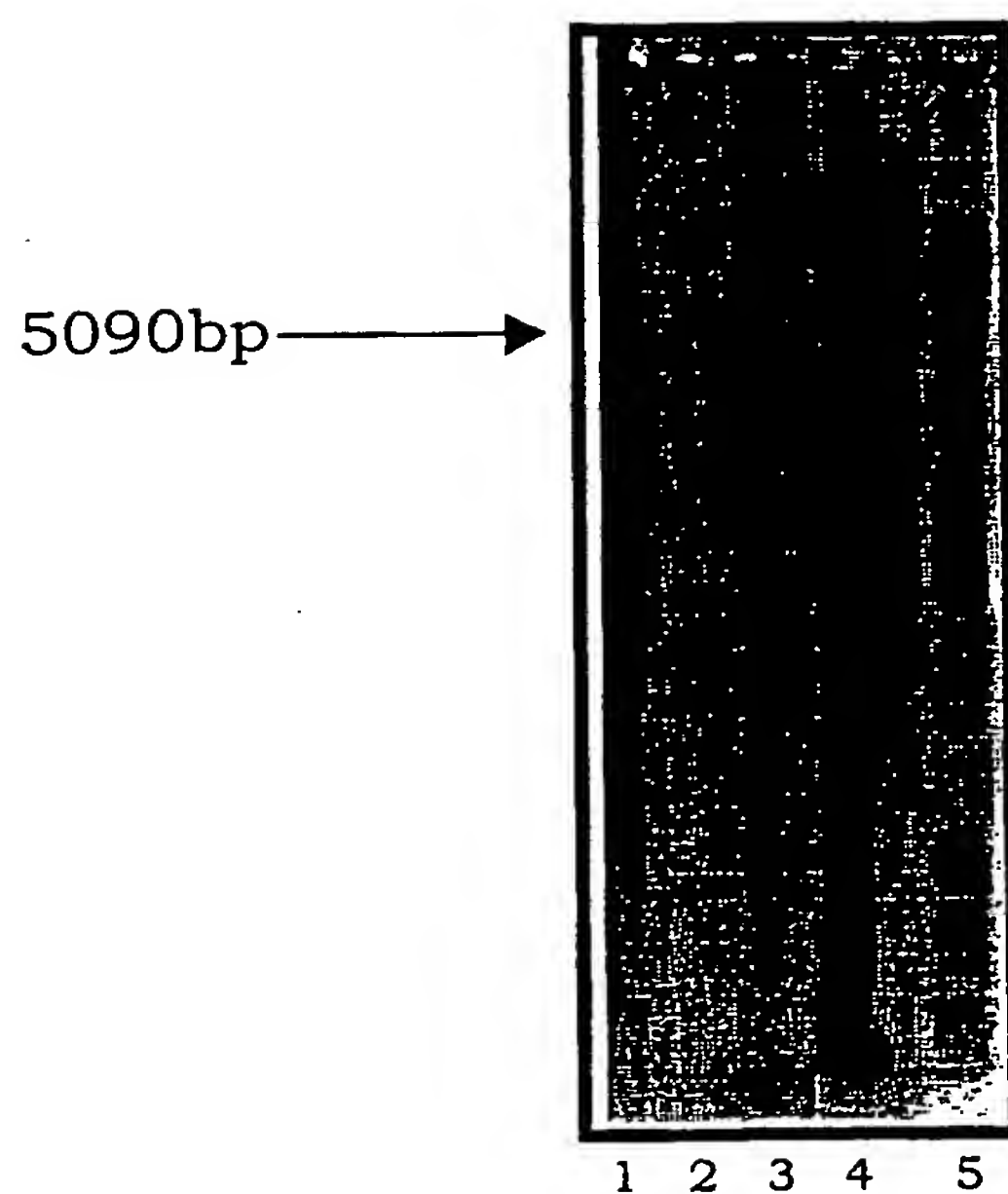


Fig. 10

13/13



## Ansatz:

	Konzentrationen	1	2	3	4
Genomische DNA	200ng/ $\mu$ l	2 $\mu$ l	2 $\mu$ l	2 $\mu$ l	2 $\mu$ l
DNTP	25mM jedes dNTP	0,8 $\mu$ l	0,8 $\mu$ l	0,8 $\mu$ l	0,8 $\mu$ l
Primer fw	20 pmol/ $\mu$ l	1 $\mu$ l	1 $\mu$ l	1 $\mu$ l	1 $\mu$ l
Primer rev	20pmol/ $\mu$ l	1 $\mu$ l	1 $\mu$ l	1 $\mu$ l	1 $\mu$ l
Puffer	100mM Tris-HCl, 500mM KCl; pH 8,3(20°C)	5 $\mu$ l	5 $\mu$ l	5 $\mu$ l	5 $\mu$ l
MgCl <sub>2</sub>	25mM	12 $\mu$ l	12 $\mu$ l	12 $\mu$ l	12 $\mu$ l
Aufbewahrungs Puffer	50%Glycerol; 100mM Tris-HCl pH 7,9, 100mM KCl; 10mM EDTA	8 $\mu$ l	7 $\mu$ l	0 $\mu$ l	8 $\mu$ l
Af PCNA	0,3 $\mu$ g/ $\mu$ l	0 $\mu$ l	1 $\mu$ l	8 $\mu$ l	0 $\mu$ l
Taqfusionprotein	Spur 4: Taq DNA Polymerase	3 $\mu$ l	3 $\mu$ l	3 $\mu$ l	TAQ 2 $\mu$ l
Wasser		17,2 $\mu$ l	17,2 $\mu$ l	17,2 $\mu$ l	18,2 $\mu$ l

## SEQUENCE LISTING

&lt;110&gt; LION Bioscience AG

&lt;120&gt; Chimäre Proteine

&lt;130&gt; LM1527PCT

&lt;160&gt; 17

&lt;170&gt; PatentIn version 3.0

&lt;210&gt; 1

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Konsensussequenz

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (2)..(3)

&lt;223&gt; Xaa steht für G, A, V, L, I, M, P, F, oder W

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (16)..(16)

&lt;223&gt; Xaa steht für G, A, V, L, I, M, P, F, oder W

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (4)..(15)

&lt;223&gt; Xaa steht für eine beliebige Aminosäure

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (17)..(20)

&lt;223&gt; Xaa steht für eine beliebige Aminosäure

&lt;400&gt; 1

Asp	Xaa	Xaa	Xaa	Xaa	Tyr	Asn	Xaa	Xaa	Xaa	Phe	Asp	Xaa	Pro	Tyr	Xaa
1				5					10					15	

Xaa Xaa Arg Ala

20

&lt;210&gt; 2

&lt;211&gt; 26

&lt;212&gt; PRT

&lt;213&gt; Konsensussequenz

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (1)..(13)

&lt;223&gt; Xaa steht für G, A, V, L, I, M, P, F, W

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (18)..(18)

&lt;223&gt; Xaa steht für G, A, V, L, I, M, P, F, W

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (15)..(15)

&lt;223&gt; Xaa steht für eine beliebige Aminosäure

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (23)..(23)

&lt;223&gt; Xaa steht für eine beliebige Aminosäure

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (25)..(25)

&lt;223&gt; Xaa steht für K, R, H, D, oder E

&lt;400&gt; 2

Ala	Xaa	Arg	Thr	Ala	Xaa	Ala	Xaa	Xaa	Thr	Glu	Gly	Xaa	Val	Xaa	Ala
1				5					10				15		

Pro	Xaa	Glu	Gly	Ile	Ala	Xaa	Val	Xaa	Ile
			20					25	

&lt;210&gt; 3

&lt;211&gt; 7

&lt;212&gt; PRT

<213> Konsensussequenz  
<220>  
<221> VARIANT  
<222> (1)..(1)  
<223> Xaa steht für M, A, L, oder M  
<220>  
<221> VARIANT  
<222> (2)..(2)  
<223> Xaa steht für Q, L, V, oder H  
<220>  
<221> VARIANT  
<222> (6)..(6)  
<223> Xaa steht für S, oder E  
<220>  
<221> VARIANT  
<222> (7)..(7)  
<223> Xaa steht für A, L, V, oder I  
<400> 3  
Xaa Xaa Val His Asp Xaa Xaa  
1 5  
<210> 4  
<211> 24  
<212> PRT  
<213> Konsensussequenz  
<220>  
<221> VARIANT  
<222> (1)..(1)  
<223> Xaa steht für G, A, V, L, I, M, P, F, oder W  
<220>  
<221> VARIANT  
<222> (6)..(6)  
<223> Xaa steht für G, A, V, L, I, M, P, F, oder W



<220>

<221> VARIANT

<222> (9)..(9)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (11)..(11)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (13)..(13)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (3)..(5)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (7)..(8)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (10)..(10)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (12)..(12)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (14)..(24)

<223> Xaa steht für eine beliebige Aminosäure

<400> 4

Xaa	Asp	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5				10						15	

Xaa	Phe	Xaa	Xaa	Tyr	Xaa	Xaa	Asp
				20			

<210> 5

<211> 10

<212> PRT

<213> Konsensussequenz

<220>

<221> VARIANT

<222> (1)..(1)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (9)..(9)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (2)..(4)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (3)..(3)

<223> Xaa steht für K, R, H, D, oder E

<400> 5

Xaa	Xaa	Xaa	Xaa	Leu	Ala	Pro	Xaa	Xaa	Glu
1				5					10

<210> 6

<211> 22

<212> PRT

<213> Konsensussequenz

<220>

<221> VARIANT

<222> (1)..(1)

<223> Xaa steht für K, R, H, D, oder E

<220>

<221> VARIANT

<222> (20)..(20)

<223> Xaa steht für K, R, H, D, oder E

<220>

<221> VARIANT

<222> (2)..(3)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (5)..(15)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (19)..(19)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (21)..(21)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (4)..(4)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (16)..(16)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (18)..(18)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<400> 6

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Gly	Xaa	Ser	Xaa	Xaa	Xaa	Leu	Xaa	Thr	Xaa
1				5					10					15	

Pro	Xaa	Xaa	Xaa	Xaa	Pro
				20	

<210> 7

<211> 7

<212> PRT

<213> Konsensussequenz

<220>

<221> VARIANT

<222> (1)..(7)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<400> 7

Xaa	Xaa	Met	Pro	Xaa	Arg	Xaa
1				5		

<210> 8

<211> 9

<212> PRT

<213> Konsensussequenz

<220>

<221> VARIANT

<222> (2)..(2)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (5)..(5)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (3)..(3)

<223> Xaa steht für S, T, oder G

<220>

<221> VARIANT

<222> (6)..(6)

<223> Xaa steht für S, D, A, oder K

<220>

<221> VARIANT

<222> (7)..(7)

<223> Xaa steht für F oder W

<220>

<221> VARIANT

<222> (8)..(8)

<223> Xaa steht für F, M, oder L

<220>

<221> VARIANT

<222> (9)..(9)

<223> Xaa steht für S, A, R, N, oder K

<400> 8

Gln Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa  
1 5

<210> 9

<211> 9

<212> PRT

<213> Thermus aquaticus

<400> 9

Gln Met Ser Leu Asp Ser Phe Phe Ser  
1 5

<210> 10

<211> 7

<212> PRT

<213> Konsensussequenz

<220>

<221> VARIANT

<222> (2)..(2)

<223> Xaa steht für L, S, oder V

<220>

<221> VARIANT

<222> (4)..(4)

<223> Xaa steht für I, L, oder V

<220>

<221> VARIANT

<222> (5)..(5)

<223> Xaa steht für L, I, D, oder S

<220>

<221> VARIANT

<222> (6)..(6)

<223> Xaa steht für V oder D

<220>

<221> VARIANT

<222> (7)..(7)

<223> Xaa steht für F, N, oder D

<220>

<221> VARIANT

<222> (3)..(3)

<223> Xaa steht für C, D, E, H, K, N, Q, R, S, oder T

<400> 10

Gln Xaa Xaa Xaa Xaa Xaa Xaa  
1 5

<210> 11

<211> 6

<212> PRT

<213> Konsensussequenz



<220>

<221> VARIANT

<222> (1)..(1)

<223> Xaa steht für Q oder W

<220>

<221> VARIANT

<222> (2)..(2)

<223> Xaa steht für L oder K

<220>

<221> VARIANT

<222> (3)..(3)

<223> Xaa steht für T oder V

<220>

<221> VARIANT

<222> (4)..(4)

<223> Xaa steht für L oder F

<220>

<221> VARIANT

<222> (5)..(5)

<223> Xaa steht für D oder F

<400> 11

Xaa Xaa Xaa Xaa Xaa Phe  
1 5

<210> 12

<211> 29

<212> PRT

<213> Konsensussequenz

<220>

<221> VARIANT

<222> (1)..(2)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (4)..(4)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (17)..(17)

<223> Xaa steht für G, A, V, L, I, M, P, F, oder W

<220>

<221> VARIANT

<222> (3)..(3)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (5)..(16)

<223> Xaa steht für eine beliebige Aminosäure

<220>

<221> VARIANT

<222> (18)..(29)

<223> Xaa steht für eine beliebige Aminosäure

<400> 12

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10					15	

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Phe	Xaa	Pro	Tyr	Phe	Tyr	Xaa
				20				25				

<210> 13

<211> 22

<212> DNA

<213> Vektorenspezifische Primer

<400> 13  
agggcgtggt gcggagggcg gt

22

<210> 14

<211> 22

<212> DNA

<213> Vektorenspezifische Primer

<400> 14

tcgagcggcc gcccgggcag gt

22

<210> 15

<211> 23

<212> DNA

<213> Vektorenspezifische Primer

<400> 15

aggaacaaca tatgacgcac tct

23

<210> 16

<211> 23

<212> DNA

<213> Vektorenspezifische Primer

<400> 16

taggtggcct gcagtaatgt tag

23

<210> 17

<211> 888

<212> PRT

<213> Thermus aquaticus und Archaeoglobus fulgidus

<400> 17

Met	Arg	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu	Leu
1				5					10					15	

Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys	Gly
			20					25					30		

Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe	Ala
		35					40					45			

Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile	Val
	50					55					60				

Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly	Gly
65					70					75					80

Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln	Leu
				85					90					95	

Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu	Glu
			100					105					110		

Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys	Lys
		115					120					125			

Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys	Asp
						135					140				

Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu	Gly	145	150	155	160
Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	Pro	165	170	175	
Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	Asn	180	185	190	
Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu	Leu	195	200	205	
Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg	Leu	210	215	220	
Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu	Lys	225	230	235	240
Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	Val	245	250	255	
Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	Phe	260	265	270	
Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	Leu	275	280	285	
Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	Gly	290	295	300	
Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	Asp	305	310	315	320
Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	Pro	325	330	335	
Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu	Leu	340	345	350	
Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	Pro	355	360	365	
Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	Asn	370	375	380	
Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	Glu	385	390	395	400
Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	Leu	405	410	415	
Trp	Gly	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	Glu	420	425	430	
Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	Gly	435	440	445	
Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	Ala	450	455	460	
Glu	Glu	Ile	Ala	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly	His	465	470	475	480

Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp  
 485 490 495  
 Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg  
 500 505 510  
 Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile  
 515 520 525  
 Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr  
 530 535 540  
 Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu  
 545 550 555 560  
 His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser  
 565 570 575  
 Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln  
 580 585 590  
 Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala  
 595 600 605  
 Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly  
 610 615 620  
 Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr  
 625 630 635 640  
 Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro  
 645 650 655  
 Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly  
 660 665 670  
 Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu  
 675 680 685  
 Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg  
 690 695 700  
 Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val  
 705 710 715 720  
 Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg  
 725 730 735  
 Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro  
 740 745 750  
 Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu  
 755 760 765  
 Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His  
 770 775 780  
 Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala  
 785 790 795 800  
 Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro  
 805 810 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu  
820 825 830

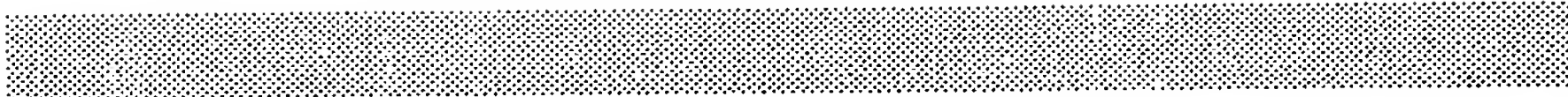
Pro Pro Arg Pro Arg Pro Lys Ser Gly Ile Glu Ile Lys Lys Leu Asp  
835 840 845

Lys Asp Tyr Tyr Ile Asp Asn Gln Ile Ile Pro Ser Val Leu Arg Ile  
850 855 860

Leu Glu Arg Phe Gly Tyr Thr Glu Ala Ser Leu Lys Gly Ser Ser Gln  
865 870 875 880

Met Ser Leu Asp Ser Phe Phe Ser  
885



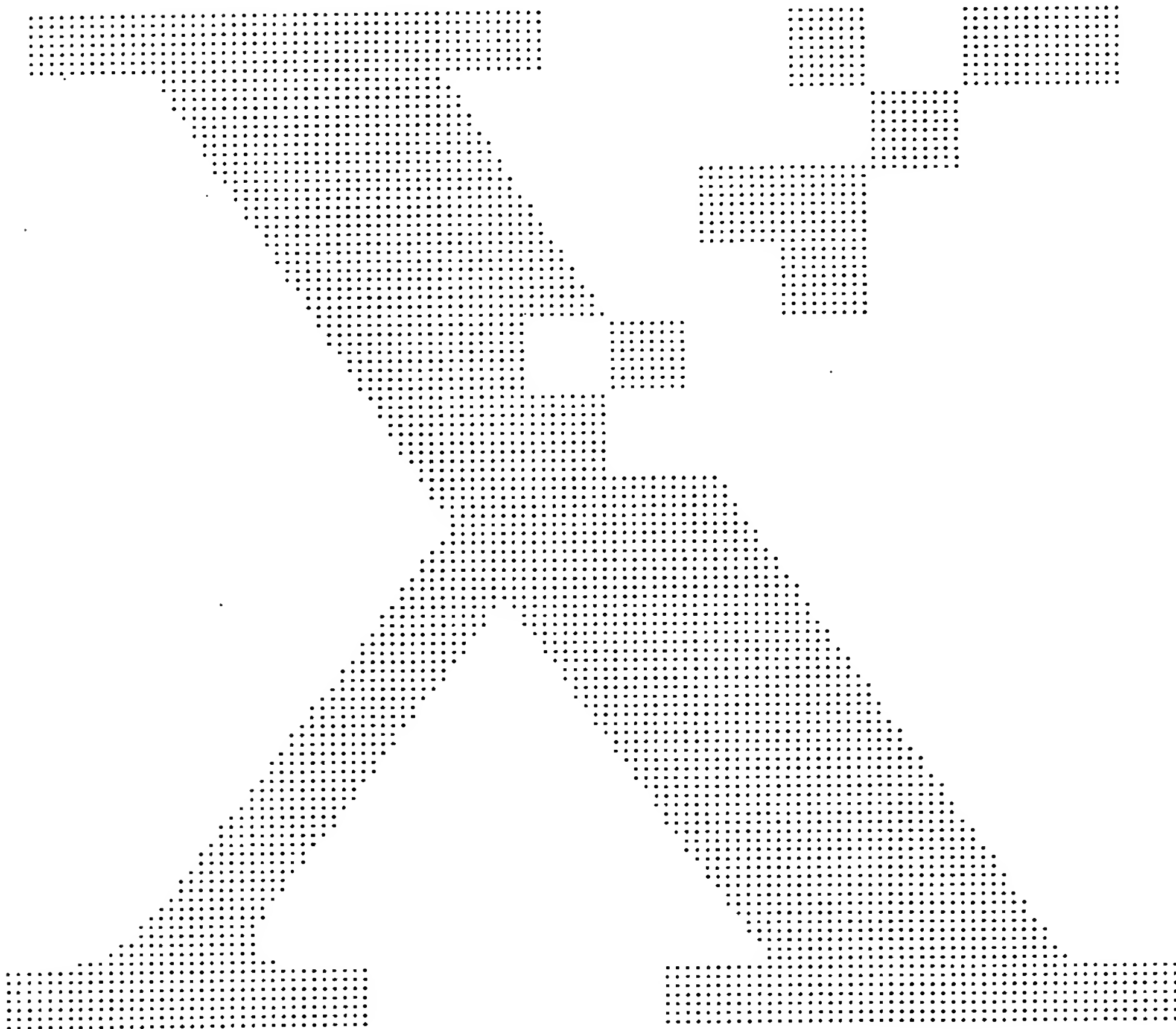


VD

# VD

 <http://www5d.delphion.com/cgi-bin/download>

 13-04-07 09:43



(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
15. Februar 2001 (15.02.2001)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 01/11051 A3

(51) Internationale Patentklassifikation<sup>7</sup>: C12N 15/31,  
15/54, 15/10, 15/62, 15/81, 9/12, 1/19, C07K 14/195,  
16/18, 16/40, C12P 19/34, C12Q 1/68

(21) Internationales Aktenzeichen: PCT/DE00/02657

(22) Internationales Anmeldedatum:  
7. August 2000 (07.08.2000)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:  
199 37 230.6 6. August 1999 (06.08.1999) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme  
von US): LION BIOSCIENCE AG [DE/DE]; Im Neuen-  
heimer Feld 517, D-69120 Heidelberg (DE).

(72) Erfinder; und  
(75) Erfinder/Anmelder (nur für US): KILGER, Christian  
[DE/DE]; Handschuhsheimer Landstr. 4, D-69121 Hei-  
delberg (DE). MOTZ, Michael [DE/DE]; Wundtstr. 9,  
D-69123 Heidelberg (DE). LÖSER, Eva [DE/DE]; Fritz-  
Frey-Str. 5/A37, D-69121 Heidelberg (DE). KÖGL, Man-  
fred [DE/DE]; Hauptstr. 131/4, D-69214 Eppelheim (DE).

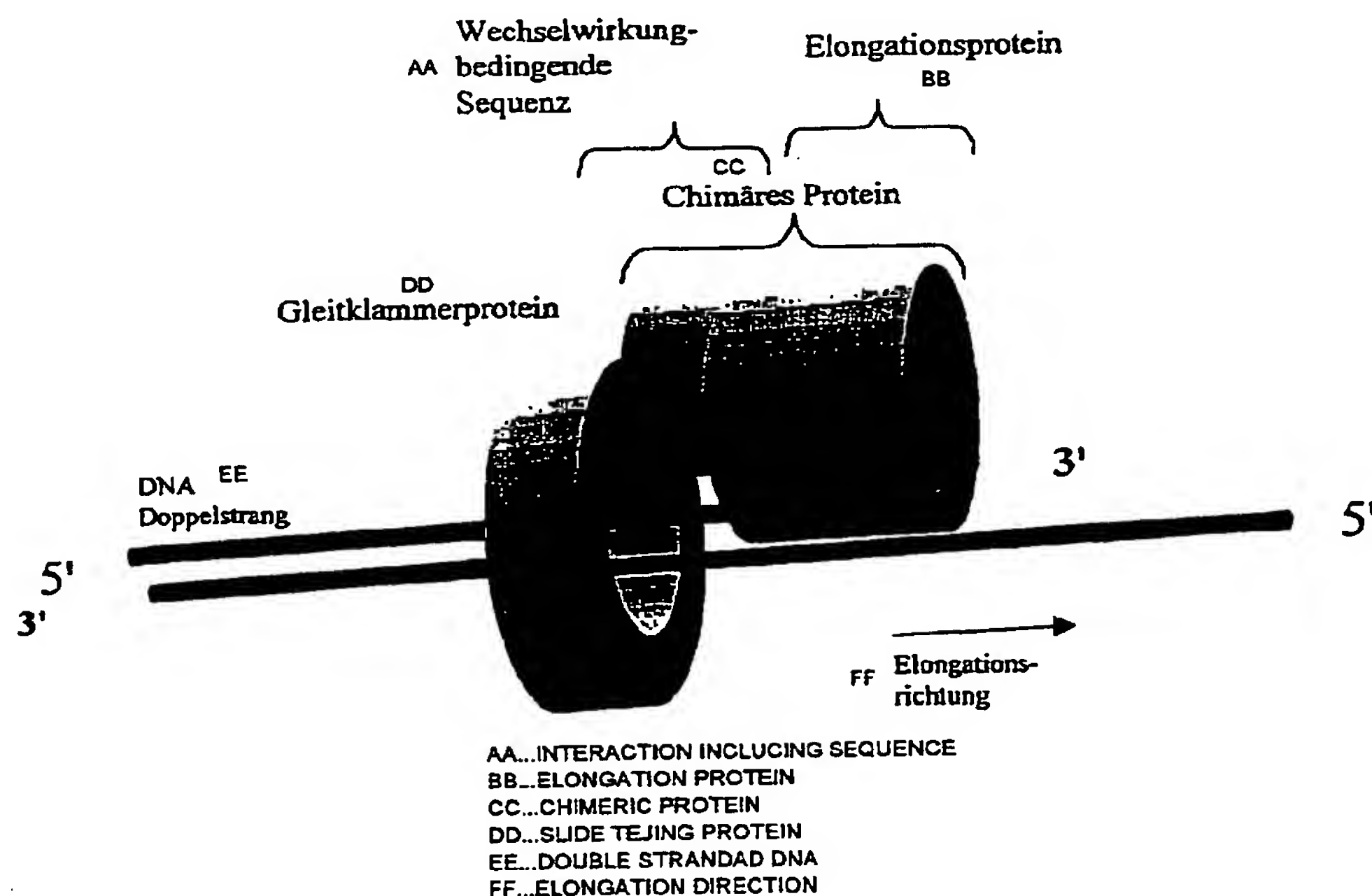
(74) Anwalt: BOHMANN, Armin, K.; Bohmann & Loosen,  
Sonnenstrasse 8, D-80331 München (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU,  
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,  
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,  
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Fortsetzung auf der nächsten Seite]

(54) Title: CHIMERIC PROTEINS

(54) Bezeichnung: CHIMÄRE PROTEINE



(57) Abstract: The invention relates to a recombinant chimeric protein comprising a) a first domain with a nucleic acid synthesis activity and b) an interaction mediating sequence, whereby said interaction mediating sequence can form a complex through the nucleic acid synthesis activity and a slide bracketing protein. Said complex is different from the complex formed through the nucleic acid synthesis activity and/or the slide tying protein with their natural interaction partner(s).

[Fortsetzung auf der nächsten Seite]

WO 01/11051 A3



(84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(88) Veröffentlichungsdatum des internationalen  
Recherchenberichts: 21. Juni 2001

*Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.*

**Veröffentlicht:**

— Mit internationalem Recherchenbericht.

(57) Zusammenfassung: Die Erfindung betrifft ein rekombinantes chimäres Protein, das a) eine erste Domäne mit Nukleinsäuresyntheseaktivität, und b) eine Wechselwirkung vermittelnde Sequenz umfaßt, wobei durch die Wechselwirkung vermittelnde Sequenz ein Komplex aus Nukleinsäuresyntheseaktivität und Gleitklammerprotein gebildet wird, der Komplex verschieden ist von dem Komplex, den die Nukleinsäuresyntheseaktivität und/or das Gleitklammerprotein mit ihren natürlichen Wechselwirkungspartner (n) ausbildet/ausbilden.

# INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/DE 00/02657

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/31 C12N15/54 C12N15/10 C12N15/62 C12N15/81  
C12N9/12 C12N1/19 C07K14/195 C07K16/18 C07K16/40  
C12P19/34 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K C12P C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, PAJ, CAB Data, STRAND, BIOSIS, EPO-Internal

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>OUHAMMOUCH MOHAMED ET AL: "Activation of RNA polymerase II by topologically linked DNA-tracking proteins."</p> <p>PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 94, no. 13, 1997, pages 6718-6723, XP002157966</p> <p>1997</p> <p>ISSN: 0027-8424</p> <p>the whole document</p> <p style="text-align: center;">--- -/--</p>	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*G\* document member of the same patent family

Date of the actual completion of the international search

19 January 2001

Date of mailing of the international search report

06/02/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

HORNIG, H

# INTERNATIONAL SEARCH REPORT

Inter national Application No  
PCT/DE 00/02657

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>PODUST VLADIMIR N ET AL: "Replication factor C disengages from proliferating cell nuclear antigen (PCNA) upon sliding clamp formation, and PCNA itself tethers DNA polymerase delta to DNA." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 48, 27 November 1998 (1998-11-27), pages 31992-31999, XP002157967 ISSN: 0021-9258 the whole document</p> <p style="text-align: center;">---</p>	
A	<p>WO 99 37661 A (UNIV ROCKEFELLER) 29 July 1999 (1999-07-29) the whole document</p> <p style="text-align: center;">---</p>	
A	<p>WO 99 13060 A (ENZYCO INC) 18 March 1999 (1999-03-18) the whole document</p> <p style="text-align: center;">---</p>	
A	<p>WO 98 45452 A (UNIV ROCKEFELLER) 15 October 1998 (1998-10-15) the whole document</p> <p style="text-align: center;">---</p>	
A	<p>WO 98 07845 A (JOUNG J KEITH ;DOVE SIMON (US); HARVARD COLLEGE (US); HOCHSCHILD A) 26 February 1998 (1998-02-26) the whole document</p> <p style="text-align: center;">---</p>	
A	<p>WO 96 37618 A (COMMISSARIAT ENERGIE ATOMIQUE ;SENTENAC ANDRE (FR); MARSOLIER MARI) 28 November 1996 (1996-11-28) the whole document</p> <p style="text-align: center;">---</p>	
P, A	<p>WO 00 08164 A (KOBEL INGO ;MOECKEL GERD (DE); VOSS HARTMUT (DE); KILGER CHRISTIAN) 17 February 2000 (2000-02-17) the whole document</p> <p style="text-align: center;">---</p>	
P, A	<p>DE 198 40 771 A (LION BIOSCIENCE AG) 10 February 2000 (2000-02-10) the whole document</p> <p style="text-align: center;">-----</p>	

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/DE 00/02657

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9937661	A	29-07-1999	AU 2341699 A EP 1056763 A	09-08-1999 06-12-2000
WO 9913060	A	18-03-1999	AU 9229698 A EP 1012248 A	29-03-1999 28-06-2000
WO 9845452	A	15-10-1998	AU 7683998 A EP 0983365 A	30-10-1998 08-03-2000
WO 9807845	A	26-02-1998	AU 4159697 A US 5925523 A	06-03-1998 20-07-1999
WO 9637618	A	28-11-1996	FR 2734567 A EP 0828838 A US 5905025 A	29-11-1996 18-03-1998 18-05-1999
WO 0008164	A	17-02-2000	DE 19840771 A AU 5617199 A	10-02-2000 28-02-2000
DE 19840771	A	10-02-2000	AU 5617199 A WO 0008164 A	28-02-2000 17-02-2000



# INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen  
PCT/DE 00/02657

<b>A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES</b> IPK 7 C12N15/31 C12N15/54 C12N15/10 C12N15/62 C12N15/81 C12N9/12 C12N1/19 C07K14/195 C07K16/18 C07K16/40 C12P19/34 C12Q1/68		
Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK		
<b>B. RECHERCHIERTE GEBIETE</b> Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole) IPK 7 C12N C07K C12P C12Q		
Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen		
Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe) WPI Data, PAJ, CAB Data, STRAND, BIOSIS, EPO-Internal		
<b>C. ALS WESENTLICH ANGESEHENE UNTERLAGEN</b>		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	OUHAMMOUCH MOHAMED ET AL: "Activation of RNA polymerase II by topologically linked DNA-tracking proteins." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, Bd. 94, Nr. 13, 1997, Seiten 6718-6723, XP002157966 1997 ISSN: 0027-8424 das ganze Dokument --- -/--	
<input checked="" type="checkbox"/> Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen		
<input checked="" type="checkbox"/> Siehe Anhang Patentfamilie		
* Besondere Kategorien von angegebenen Veröffentlichungen : *A* Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist *E* älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist *L* Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt) *O* Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht *P* Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist *T* Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist *X* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden *Y* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist *&* Veröffentlichung, die Mitglied derselben Patentfamilie ist		
Datum des Abschlusses der internationalen Recherche 19. Januar 2001		Absendedatum des internationalen Recherchenberichts 06/02/2001
Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Bevollmächtigter Bediensteter HORNIG, H

# INTERNATIONALER RECHERCHENBERICHT

Inter nales Aktenzeichen

PCT/DE 00/02657

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	<p>PODUST VLADIMIR N ET AL: "Replication factor C disengages from proliferating cell nuclear antigen (PCNA) upon sliding clamp formation, and PCNA itself tethers DNA polymerase delta to DNA." JOURNAL OF BIOLOGICAL CHEMISTRY, Bd. 273, Nr. 48, 27. November 1998 (1998-11-27), Seiten 31992-31999, XP002157967 ISSN: 0021-9258 das ganze Dokument</p> <p>---</p>	
A	<p>WO 99 37661 A (UNIV ROCKEFELLER) 29. Juli 1999 (1999-07-29) das ganze Dokument</p> <p>---</p>	
A	<p>WO 99 13060 A (ENZYCO INC) 18. März 1999 (1999-03-18) das ganze Dokument</p> <p>---</p>	
A	<p>WO 98 45452 A (UNIV ROCKEFELLER) 15. Oktober 1998 (1998-10-15) das ganze Dokument</p> <p>---</p>	
A	<p>WO 98 07845 A (JOUNG J KEITH ;DOVE SIMON (US); HARVARD COLLEGE (US); HOCHSCHILD A) 26. Februar 1998 (1998-02-26) das ganze Dokument</p> <p>---</p>	
A	<p>WO 96 37618 A (COMMISSARIAT ENERGIE ATOMIQUE ;SENTENAC ANDRE (FR); MARSOLIER MARI) 28. November 1996 (1996-11-28) das ganze Dokument</p> <p>---</p>	
P,A	<p>WO 00 08164 A (KOBBER INGO ;MOECKEL GERD (DE); VOSS HARTMUT (DE); KILGER CHRISTIAN) 17. Februar 2000 (2000-02-17) das ganze Dokument</p> <p>---</p>	
P,A	<p>DE 198 40 771 A (LION BIOSCIENCE AG) 10. Februar 2000 (2000-02-10) das ganze Dokument</p> <p>-----</p>	

# INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichung, die die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/DE 00/02657

Im Recherchenbericht angeführtes Patentedokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9937661 A	29-07-1999	AU 2341699 A EP 1056763 A	09-08-1999 06-12-2000
WO 9913060 A	18-03-1999	AU 9229698 A EP 1012248 A	29-03-1999 28-06-2000
WO 9845452 A	15-10-1998	AU 7683998 A EP 0983365 A	30-10-1998 08-03-2000
WO 9807845 A	26-02-1998	AU 4159697 A US 5925523 A	06-03-1998 20-07-1999
WO 9637618 A	28-11-1996	FR 2734567 A EP 0828838 A US 5905025 A	29-11-1996 18-03-1998 18-05-1999
WO 0008164 A	17-02-2000	DE 19840771 A AU 5617199 A	10-02-2000 28-02-2000
DE 19840771 A	10-02-2000	AU 5617199 A WO 0008164 A	28-02-2000 17-02-2000

## Elucidation of an Archaeal Replication Protein Network to Generate Enhanced PCR Enzymes\*

Received for publication, August 14, 2001, and in revised form, December 21, 2001  
Published, JBC Papers in Press, January 22, 2002, DOI 10.1074/jbc.M107793200

Michael Motz†§, Ingo Kober§¶, Charles Girardot¶, Eva Loeser¶, Ulrike Bauer‡, Michael Alberst‡, Gerd Moeckell‡, Eric Minch¶, Hartmut Voss‡, Christian Kilger‡, and Manfred Koegl¶\*\*

From the ‡Exploratory Research, ¶Proteins and Assays, and ¶Life Science Informatics, LION Bioscience Ktiengesellschaft, D-69120 Heidelberg, Germany

Thermostable DNA polymerases are an important tool in molecular biology. To exploit the archaeal repertoire of proteins involved in DNA replication for use in PCR, we elucidated the network of proteins implicated in this process in *Archaeoglobus fulgidus*. To this end, we performed extensive yeast two-hybrid screens using putative archaeal replication factors as starting points. This approach yielded a protein network involving 30 proteins potentially implicated in archaeal DNA replication including several novel factors. Based on these results, we were able to improve PCR reactions catalyzed by archaeal DNA polymerases by supplementing the reaction with predicted polymerase co-factors. In this approach we concentrated on the archaeal proliferating cell nuclear antigen (PCNA) homologue. This protein is known to encircle DNA as a ring in eukaryotes, tethering other proteins to DNA. Indeed, addition of *A. fulgidus* PCNA resulted in marked stimulation of PCR product generation. The PCNA-binding domain was determined, and a hybrid DNA polymerase was constructed by grafting this domain onto the classical PCR enzyme from *Thermus aquaticus*, *Taq* DNA polymerase. Addition of PCNA to PCR reactions catalyzed by the fusion protein greatly stimulated product generation, most likely by tethering the enzyme to DNA. This sliding clamp-induced increase of PCR performance implies a promising novel micromechanical principle for the development of PCR enzymes with enhanced processivity.

For all forms of life, the process of DNA-replication is essential in the propagation of genetic information. A complex multiprotein machinery including DNA polymerases, processivity factors, proof-reading, repair, and regulatory activities (1–3) has evolved to handle the tasks associated with this process. The importance of replication is demonstrated by the fact that its central features are highly conserved among all cellular organisms, while the protein sequences of some of the factors are not. This is exemplified by the processivity factors of DNA polymerases. Processivity is defined as the number of polymerization events during a single contact between polymerase and template. To prevent dissociation off the template DNA, many polymerases are bound by a protein that encircles the

DNA in a ring-like structure. Together with the polymerase, the ring appears to move along the DNA as replication proceeds. Such “sliding clamps” exist both for eubacteria (the  $\beta$ -subunit of the polymerase) and for eukaryotes (PCNA),<sup>1</sup> and the structures of these proteins are almost superimposable (4). However, the proteins are highly deviant in primary sequence. This situation of sequence divergence and functional conservation is also evident for the proteins responsible for loading the sliding clamp onto the DNA. Whereas in eukaryotes the clamp loader is a heteropentameric complex called RFC (2, 5), the eubacterial clamp loader is represented by the pentameric so called  $\gamma$ -complex (6, 7). Until recently very little was known about replication in the third domain of life, the archaeobacteria. In part, this is probably due to the difficult culture conditions for these often extremophile organisms. Only with the recent availability of the genomes of several archaea (8) have these organisms become amenable for bioinformatic and functional genomic analyses.

Such analyses support the notion that archaeobacteria, although possessing metabolic features that are quite similar to bacterial processes, are more closely related to eukaryotes when translation, transcription, and replication proteins are compared. For example a clear homologue of PCNA can be found in all published archaeobacterial genomes, (9) but no homologues for the eubacterial  $\beta$ -clamp or other eubacterial replication factors have been discovered. Furthermore, *bona fide* homologues of the eukaryotic DNA polymerase  $\delta$  can be identified in archaea (10).

In other cases, the situation is not as straightforward. For the eukaryotic five-subunit clamp loader RFC, only two homologues could be identified so far in archaea (10, 11). Furthermore, relatively recent work detected a completely novel two-subunit DNA polymerase (DP1 and DP2) with assumed replicative function in thermophilic archaea, the large subunit of which displays no significant homology to any other DNA polymerase (12, 13).

Due to their use in DNA-sequencing and PCR applications, heat-stable DNA polymerases are technically and economically important enzymes. Current PCR reactions rely on single enzymes, which like the DNA polymerase from *Thermus aquaticus* (*Taq*) often originate from cellular DNA repair polymerases (14). Others represent potential replicative enzymes such as *Pfu* and *Pwo* (15), but are employed in the absence of any accessory replication factor. For longer PCR templates, artificial enzyme mixes with compromised fidelity are the state of the art (16). Remarkably, however, these reactions are orders

\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Both authors contributed equally to this work.

\*\* To whom correspondence should be addressed: LION bioscience AG, Im Neuenheimer Feld 515-519, D-69120 Heidelberg, Germany. Tel.: 49-6221-40-38-140; Fax: 49-6221-40-38-401; E-mail: manfred.koegl@lionbioscience.com.

<sup>1</sup> The abbreviations used are: PCNA, proliferating cell nuclear antigen; RFC, replication factor C; Y2H, yeast two-hybrid; S, small fraction; L, large fraction; YPDA, yeast extract/peptone/dextrose with adenine; SD-HUL, synthetic medium-histidine, uracil, and leucine.



of magnitude less efficient than the *in vivo* DNA replication process with respect to fidelity, speed, and product size. The replacement of current monomeric PCR enzymes by a selected subset of proteins taken from the replication machinery of a thermophilic archaeobacterium may hence open new chances to improve PCR.

In this report a biochemical, bioinformatic, and functional proteomic approach was undertaken to deepen our understanding of archaeal replication to use this knowledge for the improvement of PCR applications. We demonstrate that processivity factors can enhance PCR catalyzed by archaeal polymerases. Finally, we show that by tethering *Taq* polymerase to a sliding clamp via a PCNA-interaction domain one can generate PCR enzymes with increased performance.

#### EXPERIMENTAL PROCEDURES

**Bioinformatic Analysis of Replication Proteins in *Archaeoglobus fulgidus***—In the first step of the bioinformatics analysis most of the key components of the DNA replication machinery in humans were identified. The following sequences were retrieved from the SWISSPROT database (26): AC11\_HUMAN, AC12\_HUMAN, AC13\_HUMAN, AC14\_HUMAN, AC15\_HUMAN, PCNA\_HUMAN, DPD2\_HUMAN, and DPOD\_HUMAN. In a first step, these sequences were compared with BLASTP (27) against a non redundant database of all publicly available protein sequences. In a next step, homologous sequences from related organisms including *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* from the blast result were used to generate a multiple-sequence alignment using the ClustalW program (28). Based on the obtained alignment a Hidden Markov Model was generated with the HMMER software package (version 1.8, Sean Eddy, Dept. of Genetics, Washington University School of Medicine, St. Louis, MO). This model includes the highly conserved regions but also the less homologous parts of the sequence alignment in between. Now the non-redundant sequence database was again searched with the HMMER program for distantly related homologues especially in *A. fulgidus*. Based on the sequence information of TIGR (8), a database of all *A. fulgidus* sequences was generated. This database was also searched in parallel. The following sequences were identified: AF2060 homologous to AC11\_HUMAN, AC12\_HUMAN, AC13\_HUMAN, and AC14\_HUMAN; AF1195 homologous to AC15\_HUMAN; AF0335 homologous to PCNA\_HUMAN; AF1790 homologous to DPD2\_HUMAN; and AF0497 homologous to DPOD\_HUMAN. Another protein with polymerase activity was identified by applying the same methods starting from a recently published sequence in *P. furiosus*: DP2L\_PYRFU (12). This protein is homologous to AF1722.

Using these strategies we retrieved sequences from *A. fulgidus*, which are likely part of the replication machinery. The analysis was performed using the described bioinformatics tools, which were implemented in bioSCOUT (LION bioscience AG, Heidelberg, Germany, lionbioscience.com/solutions/bioscout), a multifunctional sequence analysis program package. bioSCOUT also includes all publicly available databases as well as proprietary and in-house databases. All sequence analysis runs were made on LION's in-house servers on database content of about one terabyte. An automatic alert service implemented in bioSCOUT was used to retrieve all newly entered sequences in all major sequence databases (e.g. NCBI, EMBL, SWISSPROT, PIR), related to the described project.

**Two-hybrid Methods**—For the construction of a genomic library of the 2,178,400-bp sized *A. fulgidus* genome for yeast two-hybrid (Y2H) screening, genomic DNA was fragmented by sonication and cloned into pGAD424; 16 µg of genomic DNA of *A. fulgidus* (obtained from the German Collection of Microorganisms and Cell Cultures) were sonicated for 8 s, such that fragments ranging from 0.1 to 5 kb were generated. To minimize the cloning bias, the ligations and transformations were done in two size classes in separate reactions. Size classes were 0.3–0.7 kb (small, the S fraction) and 0.7–2.5 kb (large, the L fraction). DNA fragments of the two size classes were isolated by cutting out the respective areas of a 1% agarose gel after electrophoretic isolation in TBE. 300 ng and 450 ng of sonicated DNA of the S and L fractions, respectively, were filled in with 2.5 units of *Pwo* polymerase (Roche Molecular Biochemical) and 35 µM of an equimolar dNTP mix (Roche Molecular Biochemical) for 30 min at 72 °C in the buffer provided with the *Pwo* polymerase. After the reaction, the DNA was purified using the QIAquick PCR purification kit (Qiagen). 5 µg of pGAD424 were digested with *Sma*I (New England Biolabs) and dephosphorylated

with calf intestine phosphatase, (Roche Molecular Biochemical) and purified from an agarose gel after electrophoresis. The DNA fragments were ligated into the linearized vector at 16 °C overnight using 5 units of T4 DNA ligase (United States Biochemicals) each time in 50 µl using the buffer provided by the manufacturer and a total amount of DNA of 170 or 300 ng for the S and L fraction, respectively. The optimal ratio of vector to insert mass had been determined beforehand and was 1:1.5 for the small fraction, and 1:2 for the large fraction. The library was transformed into DH10B *Escherichia coli* cells by electroporation, yielding  $2 \times 10^5$  and  $3.5 \times 10^6$  independent colonies for the S and L fractions, respectively. Colonies were washed off, DNA was prepared, and the library was transformed into haploid yeast PJ69–4α (25) using 200 and 300 15-cm dishes for the small and large insert fractions, each plate containing 1000–2000 colonies. Each plate was then washed off with YPDA containing 20% glycerol and stored separately in the wells of 2-ml deep well microtiterplates at –80 °C. Thus, one copy of the library was contained in a set of five microtiterplates. As working stocks, the library was diluted to an optical density at 600 nm ( $OD_{600}$ ) of 5.0. The final library included roughly 500,000 independent clones. Thus, fusion proteins will be generated every 4.4 bases, resulting in an in-frame fusion every nine amino acids, on average.

For the generation of a library of fragments of the AF0497/PolB gene, we followed the same protocol as for the generation of the library with the following alterations. We subjected a PCR fragment representing the coding region of the gene to sonication for 20 s. Fragment size was 0.05–0.3 kb. Blunt fragment ends were generated using *Pwo* polymerase as described above, and fragments were cloned into *Sma*I-linearized pGBDU and pGAD424.  $10^4$  independent colonies were collected for both vectors after transformation in *E. coli*. DNA was prepared and transformed into PJ69–4α and PJ69–4α, to yield  $5 \times 10^5$  yeast colonies. Cells were washed off the plates as described above and frozen. These libraries were used to isolate clones coding for protein fragments interacting with AF0335/PCNA by mating of yeast cells as described below.

Bait and prey vectors were generated by the two-step PCR protocol as described in Hudson *et al.* (17). 5' extension for the 2nd PCR was GAATTCGGTACCACCACCATG, and 3' extension was GATCCCCGGGGAATTGCCATGTC.

For the mating of yeast cells, bait proteins were cultured overnight to an  $OD_{600}$  of 1–2 in selective medium. For each bait construct, four independent clones were cultured and mixed in equal amounts of cells, to minimize the chances of working with a clone carrying detrimental mutations from the PCR reactions. Library cells were thawed out and allowed to recover for 1 h at ambient temperature. For screens, an 0.5-OD solution of the bait mix was prepared in YPDA. 10 µl of the library were transferred from the library plates to new 96 deep well plates that contained 10 µl of YPDA and allowed to recover for 2 h at 30 °C. After the incubation, 90 µl of the bait solution were added. Mating was allowed overnight, shaking at 1000 rpm in a microtiterplate shaker. The mating mixture was then diluted with 1.1 ml of SD-HUL containing varying amounts of 3-amino triazole and cultivated for 6–10 days at 30 °C.

For the pairwise mating, the bait mix and the prey mix were prepared at a concentration of 1 OD/ml in SD-HUL and SD-L, respectively. 50 µl of each were mixed together for the mating. The mating mixture was then diluted with 1.1 ml of SD-HUL containing varying amounts of 3-amino triazole and cultivated for 6–10 days at 30 °C.

However, mating efficiencies were not satisfying in microtiterplates, and mating in flasks proved to be the superior method in our hands. Thus, for the second round of library screens, the small and large fractions of the library were combined into two separate pools, and mating was done in 50-ml Erlenmeyer flasks in YPDA at an OD of 1.0 in a total volume of 10 ml, shaking for 5 h at 100 rpm. In this case, after the mating cells were washed in selective medium lacking histidine, uracil, and leucine (SD-HUL), and aliquoted into microtiterplates in the same medium.

For both protocols as well as for library screens, after 6–10 days the cells were passaged once to microtiterplates containing fresh selection medium and allowed to grow for 2–4 days. Potential positives were identified from wells displaying cell growth. The activation of the secondary reporter genes *ADE2*, *lacZ*, and *Mel1* was determined, and double-positive cells were collected. For ease of measurement, activation of *Mel1* was determined using 5-bromo-4-chloro-3-indolyl-α-D-galactopyranoside (X-α-Gal, CLONTECH) and scoring for blue color. The inserts were amplified by PCR from the yeast cells, sequenced, and analyzed using standard bioinformatic methods as well as bioSCOUT for annotation. Interaction networks were visualized using the automated viewer piSCOUT<sup>®</sup> and reviewed for sticky proteins identified with multiple unrelated baits.